

Torlódások vizsgálata számítógépes rendszerekben a statisztikus fizika módszereivel

Stéger József

Eötvös Loránd Tudományegyetem
Természettudományi Kar
Fizika Doktori Iskola

Statisztikus fizika, biológiai fizika
és kvantumrendszerek fizikája program

Iskolavezető: Dr. Horváth Zsolt, az MTA rendes tagja

Témavezető: Dr. Vattay Gábor, egyetemi tanár

Komplex Rendszerek Fizikája Tanszék

Doktori értekezés



2009.

Tartalomjegyzék

1. Bevezetés	3
2. Fogalmak	9
2.1. A hálózati elemek fizikai paraméterei	10
2.2. A kiszolgálási sorokról	10
2.3. Az egyedi csomagküldési dinamikáról	11
2.4. Szimulációs környezetek	12
3. A mérőhálózat	15
3.1. Motiváció	15
3.2. Helyi monitorozó mérések	15
3.3. Végpont-végpont mérési módszerek	16
3.3.1. Passzív mérési eljárások	17
3.3.2. Aktív mérési eljárások	18
3.3.2.1. Egy útvonal feltárása a hálózatban	19
3.3.2.2. A hálózati késleltetés mérése	23
3.4. Az Etomic mérőhálózat	25
3.5. A Dimes program és a PlanetLab hálózat	29
3.6. Diszkusszió	34
4. Hálózatomérési virtuális obszervatórium	35
4.1. Motiváció	35
4.2. Elérhető referencia nyersadatmegosztók	36
4.3. Az obszervatórium prototípusa	37
4.4. Diszkusszió	41
5. Hálózattomográfia	43
5.1. Motiváció	43
5.2. Y-tomográfia	46
5.2.1. A becslési módszer	46
5.2.1.1. Tetszőleges fa lefedése Y-okkal	51
5.2.2. Mérési eredményeink	52
5.2.2.1. Analízis	54

5.2.2.2.	Az Internet forgalom egy napja	57
5.3.	Tomográfiás sorhosszeloszlás-beclés fán	59
5.3.1.	A beclési módszer	60
5.3.2.	Mérési eredményeink	64
5.3.2.1.	Az algoritmus futásidejének vizsgálata	65
5.3.2.2.	Analízis	68
5.4.	A traceroute-tomográfia ötlete	73
5.5.	Diszkusszió	73
6.	Torlódási hullám modellje	75
6.1.	Motiváció	75
6.2.	A gyűrű alakú hálózat	76
6.3.	Szimulációs eredmények	78
6.4.	A hullámfront kialakulása	81
6.4.1.	A hullámfront burkolója	82
6.5.	A dugóhullám stabilitása	84
6.6.	Diszkusszió	85
7.	Összefoglalás	87
8.	Summary	89

1. fejezet

Bevezetés

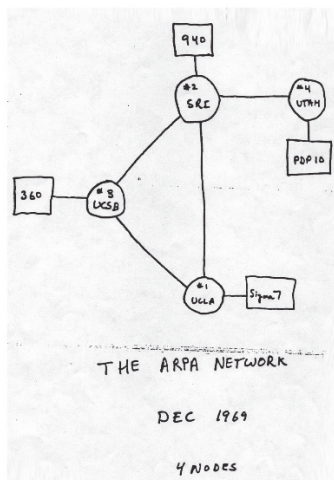
Talán a legfontosabb kérdések, melyek a hétköznapi emberekben és a kutatókban egyaránt felmerülnek a tudomány bármely területén folytatott kutatás megkezdése előtt, azok a következők: *„Miért hasznos az, amivel foglalkozunk? Mire lehet a gyakorlati életben használni a kutatás eredményeit?”* Úgy érzem, ezek a sarkított kérdések mozgatják előre a mindennapi munkánkat. Ezeket szem előtt tartva, a rájuk adható válaszainkat folyamatosan finomítva érdemes a kutatómunkát végezni. Munkánk során igyekszünk sosem szem elől tévesztetni ezt a fonalat, és minden elért lényeges mérföldkőnél újra elővesszük és értelmezzük ezeket a kérdéseket. A dolgozat főbb fejezeteinek a Motiváció és a Diszkusszió részeiben e kérdésekre kerestük a munkát érintő válaszokat.

A jelenkor kutatói előtt megoldásra váró problémák jelentős része nagyon összetett. Legtöbb esetben nem is nyílik arra lehetőség, hogy a tárgyalás végén egzakt formulákkal leírt, de a valóságot hűen leíró modell szülessen. A kutatás tárgyát képező makroszkopikus rendszert egészében vizsgálva számtalan, egyenként is több paraméterrel leírható, mikroszkopikus részrendszerre bontható. Még ha jól ismerjük is a részrendszereket egyenként leíró mozgásegyenleteket, továbbá alkalmas modell áll rendelkezésünkre a kölcsönhatások tárgyalására, a teljes rendszer viselkedéséről számot adó állapotleírók nehezen megfoghatóak maradnak. A feladat bonyolultságát tovább tetézi, ha a kutatott rendszerben véletlen folyamatok is munkálnak. Ilyen esetben a termodinamika és a statisztikus fizika eszköztárához kell fordulnunk, hogy sikeres leírást adhassunk a komplex rendszerekről.

Több izgalmas példát is felsorolhatunk azokra a hétköznapi komplex rendszerekre, amikkel a fizikus szakma érdemben foglalkozik. Ilyenek többek között a tőzsdén tapasztalható különböző időskálán megfigyelt árfolyammozgások, melyek mögött egyedi adás-vétel a mozgató rugó; az állatcsordák mozgásában megfigyelhető mintázatok; a kritikus szinkronizációk és azok elhalása egy színházi előadás végén hallható tapsviharban; a járványterjedés; a biológiai rendszerek csatolt reakciói és a szervezet együttes nyersanyaganyagháztartása. . .

A munkánk során vizsgált rendszer a számítógépes hálózatok, illetve ennek mindennapi életünket átszövő még magasabb szintű megvalósulása, az Internet

is egy komplex rendszer. Nevének igazi jelentése „*hálózatok hálózata*”,¹ ami jól sugallja, hogy ez az egész Földet beborító hatalmas – folyamatosan növekvő – rendszer, valójában egyedi számítógép-hálózatokat fog össze. Az Egyesült Államok Védelmi Minisztériumának (DARPA²) 1958-ban alapított kutatásokért felelős részlege támogatta azt a kutatócsoportot, melynek tagjai közel 40 évvel ezelőtt, 1969-ben létrehozták a mai Internet őseit, az ArpaNet-et. Az ArpaNet kezdetben négy számítógép központból épült fel, és a hálózatra megtervezett forgalomvezérlés alapozta meg az azóta is uralkodó információáramlási dinamikát leíró protollokat. Az 1.1. ábrán vázolt történelmi jelentőségű ArpaNet hálózat



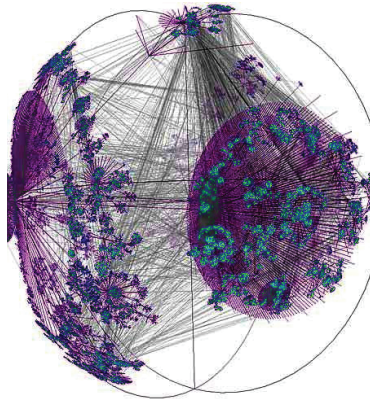
1.1. ábra. Az ArpaNet vázlata Dr. Leonard Kleinrock 1969-beli jegyzetei alapján. A mai Internet őseinek tekinthető ArpaNet hálózat négy gépének otthont adó intézmények sorában szerepel a Stanford Research Institute (SRI) és a Kaliforniai Egyetem (UCLA). A számítógépeket összekötő hálózat az AT&T telefontársaság vonalait bérelte a földrajzi távolság leküzdésére. A kép forrása: [1].

topológiája még könnyen átlátható. Mára azonban az Internet méretét 600-millió végpontra becsülik, azaz mérete eddigi élete során nyolc nagyságrenddel nagyobb lett. A hálózat topológiájának bonyolultságát az 1.2. ábrán szemléltetjük.

Az Internet valójában nem egy konkrét fizikai hálózat, hanem egy módszer

¹Az Internet szó az angol *internetwork*-ból ered.

²Defense Advanced Research Projects Agency



1.2. ábra. Korunk Internetének topológiáját ábrázoló gráf. Jól látható, hogy az évek során az összeköttetések komplexitása jelentősen megnőtt. Az ábra forrása: [2].

arra, hogy egymástól különböző, akár földrajzilag is elválasztott hálózatokat kössön egybe úgy, hogy azok egymással kommunikálni tudjanak. Így nem is létezik olyan központi szerv, ami az Internet egészének fejlődéséért egy személyben felelőse lenne. Az Internetbe kapcsolt egyedi alhálózatok rendelkeznek önálló felügylettel, saját szabályaik szerint üzemelnek és önálló életet élnek. Az alhálózatok csatlakoztatásának összehangolását, illetve az ezzel kapcsolatosan felmerülő mérnöki tevékenységeket az 1992. januárjában létrehozott, profitmentes Internet Society (ISOC) irányítja, melynek központja az amerikai Virginia állambeli Restonban székel. A társaság magyarországi képviselője a Magyar Internet Társaság Egyesület. Az Internet működtetéséért felelős regulák tehát heterogén működési környezetet biztosítanak, a különböző alhálózatok belső struktúrája folyamatosan változhat a többi alhálózat tudta nélkül: gépeket állíthatnak üzembe, vonalakat bővíthetnek, szüntethetnek meg, módosíthatják a megengedett forgalom fajtájának spektrumát.

Az Internet, építőköveit vizsgálva, sem tekinthető egyöntetű rendszernek. Az egyedi hálózatokat felépítő eszközök és az azokat összekötő vonalak nagyon széles skáláját ismerjük, és ezek folyamatosan bővülnek az új technológiai megoldásokkal. A hálózatépítésre jellemző egyik fontos vezérlő elv, hogy más kommunikációs hálózatokból üzemképpen kapott infrastruktúrákat és módszereket is beleolvasztanak a teljes rendszerbe, így például a telefonvonalak vagy a műholdas összeköttetések is részét képezik az Internetnek. Az összekötő csatornák heterogenitását

jól jellemzi a következő felsorolás, a teljesség igénye nélkül: léteznek műholdas, rádiós, optikai szálon vagy különböző típusú rézből készített vezetékeken kialakított vonalak. Az említett vonalakon különböző mértékű zaj rakódik rá a hasznos jelre, különböző mértékben fordulhatnak elő hibák. A számítógépes hálózatok kommunikációja során előforduló hibák könnyebb javításának érdekében feladták az addig alkalmazott vonalkapcsolt kommunikációs elvet, és a közvetíteni kívánt információt csomagokra bontva, a hálózati elemeken közlekedve, helyi döntések alapján irányítják célba. Fontos megemlíteni, hogy az alhálózatok között található csomópontok illesztésénél támasztott követelményeket, mindenki számára elérhető szabványrendszerben³ rögzítik. A különböző információs közegre egyedileg jellemző adattovábbítás úgy valósul meg, hogy a hálózat végpontjain működő alkalmazások számára praktikusán láthatatlan, milyen vonalon zajlik a kommunikáció.

Egykor az Internet kizárólag katonai vagy oktatási intézmények és a kutatók számára volt elérhető, manapság azonban a mindennapi életünk elengedhetetlen részét képezi. A hálózat segítségével működő szolgáltatásokat felhasználjuk munkánk és szabadidős tevékenységeink szervezésében. Társadalmi és kulturális életünkre erőteljes visszahatással van, egyre jobban függünk az Internet használatától. Éppen ezért megbízható és kiszámítható működése gazdasági és társadalmi érdek.

Bár az Internet ember alkotta rendszer és az egyedi felépítő elemek működése jól ismert, az együttes viselkedés, a bonyolult hálózati geometria és az alkalmazások erőforrásokért folytatott dinamikus harca következtében nehezen tárgyalható. Egyre nagyobb az igény jó, minél szélesebb körben alkalmazható modellek megalkotására. Kezdetben az Internetet praktikus mérnöki szempontok alapján vizsgálták, ám az idő során összegyűlt mérési és a szimulációs modellezésekből származó adatok sokasága ráébresztette a kutatókat, hogy komplex rendszerként lehet és kell vizsgálni az Internetet is [3]. Nem csoda hát, hogy a természettudomány érdeklődési körébe került az Internet és a hozzá köthető jelenségek vizsgálata és modellezése. Az Internetet, mint kommunikációs hálózatot számos kutató vizsgálja [4].

Barabási és Albert [3] áttekintő munkájukban megvizsgálják az Internet topológiájára is érvényes véletlen gráf modelleket. Számba veszik azokat az építő- és rendezőelveket, amelyek az empirikus adatokkal összevetve megválaszolják a komplex gráfok legfontosabb statisztikai jellemzőiről és a dinamikai tulajdonságokról feltett kérdéseket. Faloutsos és szerzőtársai [5, 6, 7] tanulmányaikban a hálózat érdekes tulajdonságait mutatják meg, úgymint a hálózati elemek összeköttetési grájára jellemző hatványfüggvény szerint eső fokszámbeloszlás, akár az útvonalválasztó egységek szintén, akár autonóm rendszerek⁴ szintjén képezzük

³Az Internetet érintő szabványok elfogadását és rendszerezését az Nemzetközi Szabványügyi Hivatal (ISO) végzi.

⁴Az egy szabályozás alá eső (al)hálózati elemek összességének autonóm rendszer a neve.

azt. Vazquez, Pastor-Satoras és Vespignani [8, 9] munkáikban az Internet és az autonóm rendszerek topológiáját és a gráfjaik időfejlődését több évre visszamenőlegesen elérhető mérési adatok segítségével vizsgálják. Megállapítják, hogy a hálózat topológiája skálafüggetlen tulajdonságokat mutat és a növekedési folyamat során a kötési valószínűségek lineáris preferencia-függést mutatnak az élszám függvényében. Rámutatnak továbbá arra, hogy a hálózat hierarchiájából adódó struktúrák jelennek meg a klaszterezettségi együttható és a korrelációs függvények tulajdonságaiban. Az Internet kutatásával kapcsolatos szakirodalomban a hálózat topológiájára és az ezen keresztül megfogható hiba- és támadástűrésére irányuló kérdésekkel Albert és szerzőtársai [10] keresték a választ. A hálózati topológia folyamatos fejlődését, a kötések átrendeződésének vizsgálatára alkotott modellt Palla szerzőtársaival [11], amely modell a rácsgázok analógiájának tekinthető, és számot ad a gráfátrendeződéseknél tapasztalható fázisátalakulásokról.

A hálózat folyamatos növekedésének fényében, a topológia kutatási területe mindig érdekes téma marad, de a teljes rendszer leírása szempontjából szükséges a topológiai adatokat a forgalom dinamikájára jellemző tulajdonságokkal is felfértezni. Az erre irányuló úttörő lépéseket Vazquez [8] és de Menezes [12] tették meg, akik a forgalom ingadozásait jellemző skálázási tulajdonságokat tártak fel. Számos munka tűzi ki céljául, hogy a topológia és a hálózati forgalom között tárjon fel hasznos összefüggéseket. Például Tadić és szerzőtársai tanulmányaikban [13, 14] a skálafüggetlen gráfokon kialakuló forgalom tulajdonságait vizsgálják. A hálózat puffereit összekapcsoló kölcsönhatásokat tárják fel, amiért az időről-időre korrelálttá váló csomagfolyamok felelősek. Hosszú távú korrelációkat mutatnak ki a csomagok beérkezési idősorainak teljesítményspektrumában. Szimulációs eredmények alapján a várakozási sorokban eltöltött idő hatványfüggvény szerint esik, és az átlagos sorbanállási idő a hálózat méretével skáláz. Sreenivasan és szerzőtársai [15] különböző útvonalválasztási stratégia mellett vizsgálják a hálózati topológia hatását a sorbanállási időkre és a torlódások kialakulására.

Mérésekre alapozva Leland és szerzőtársai [16, 17] kimutatták, hogy az alhálózatokban megvalósuló forgalom fraktáltulajdonságokat mutat. Veres és Boda [18], hálózat szimulátorban végzett kutatásaikra alapozva, rávilágítanak arra, hogy a forgalom dinamikáját leíró szabályok inherensen hozzájárulnak a fázistérbeli kaotikus pályák kialakulásához. Csabai [19] mérései adatokon nyugvó adatokból, a csomagok körülfutási idejének spektrumát elemezve $1/f$ -zajra jellemző tulajdonságokat állapított meg, ugyancsak a forgalmat jellemző $1/f$ tulajdonságokat talált Takayasu szerzőtársaival [20] munkájában.

A fent említett érdekes kutatási területek közül, munkám során – kollégáimmal együttműködve – a hálózati forgalom vizsgálatára és mélyebb megértésére tettem elméleti és gyakorlati lépéseket. A hálózaton „beszélgető” alkalmazások pillanatnyi forgalmára jelentős hatással vannak a hálózati késleltetések, melyek számot adnak a hálózat aktuális teljesítőképességéről. A várakozásokból eredő extrém késleltetések torlódáshoz vezetnek, amik természetesen rontják a kommunikáció feltételeit, és következképp előfordulási valószínűségük minimalizálendő. Kuta-

tásunk célja, hogy képet alkossunk a számítógépes kommunikációs hálózatokat jellemző késleltetésekről és feltárjuk az ebből eredő esetleges torlódásokat.

A dolgozat a 2. fejezetében a számítógépes hálózatok történelmi áttekintését adjuk és bevezetjük a tárgyalás megértéséhez szükséges fogalmakat. A kutatómunkánk gerincét képező kísérletek elvégzéséhez és az európai Internet formai térképének megalkotásához mérőrendszert terveztünk és építettünk, melyet a 3. fejezetben ismertetünk. Ezt követően a hálózati mérések virtuális obszervatóriumának alapjait fektettük le, hiszen a hálózati mérések archiválása és publikálása fontos, a tudomány előrehaladását szolgáló feladat. Az obszervatórium ötletét röviden a 4. fejezetben mutatjuk be. A kutatócsoportunk hálózati mérései közül az egyik legfontosabb méréstípus a hálózattomográfia, melynek segítségével a hálózat belsejét leíró, közvetlenül nem mérhető késleltetések statisztikai jellemzőit becsüljük meg. Az 5. fejezetben tomográfias kísérleteink nyersadatait kiértékelő statisztikus fizikai módszereket mutatjuk be. Kísérleteink tanúsága szerint, a forgalom fraktál tulajdonságára utaló jeleket mutat. A sorbanállási időkre jellemző új skálázási tulajdonságot fedezünk fel.

A 6. fejezetben egy elméleti modellt alkotunk, amely számot ad a hálózatban kialakuló torlódási mintázat tér- és időbeli fejlődéséről. Modellünk a hálózati szimulátorok jóslataival összecsengő eredményekkel szolgál.

A 7. és a 8. fejezet a dolgozat magyar és angol összefoglalása.

2. fejezet

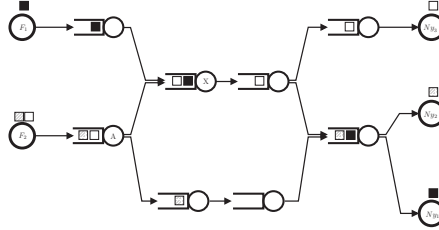
Az Internet modellezéséhez szükséges alapfogalmak

E fejezet célja, hogy tömören összefoglalja azokat az alapvető fogalmakat, amelyek a számítógépes kommunikációs hálózatok modellezésével foglalkozó irodalom és kutatások mindennapos szókincsét alkotják.

A technikai valóság részleteit leegyszerűsítve, de a lényeget megragadva, a számítógéphálózatokat alkotó aktív elemeket, két csoportra bonthatjuk működésüket tekintve. Lehetnek végpontok vagy útvonalválasztó egységek, idegen szóhasználatnál *terminálok* vagy *routerek*. Ezeket az eszközöket passzív, irányított vonalak, *linkek* kötik össze. Az esetek jelentős részében, a hálózatokon a kommunikáció végpont-párok között, irányítottan valósul meg. Az egyik végpont a küldő, a másik végpont a fogadó szerepét tölti be. Az átvitelre szánt információ feldarabolva, szabványos csomagok folyamaként közlekedik a hálózati elemeken. A folyamatot alkotó összes csomagban megtalálható a címzés, ami elengedhetetlen ahhoz, hogy a hálózat csomópontjaiban az útvonalválasztó egységekben, – azaz ahol több vonal összefut és több vonal vezet ki, – megoldható legyen a csomagok helyes irányítása. A kommunikációra jellemző, hogy a csomagok egymástól teljesen függetlenül közlekednek, útvonalválasztásuk helyi döntések alapján történik. Tehát az Interneten nem követelmény az, hogy a csomagok a kommunikáció során végig ugyanazon az úton haladjanak, amennyiben több útvonal is található a „beszélgető” végpont-pár között. Mérési tapasztalataink azonban (lásd a 3.3.2. fejezet) azt mutatják, hogy a jelenkori Interneten minden információs folyam első csomagja, mintegy kifűrja magának a lehetséges útvonalak közül azt az egyet, amelyet ezután az információs folyam többi csomagja is bejár.

Természetesen a hálózatban, lásd a 2.1. ábrán, egyidőben több végpont-pár beszélget, csomagjaik találkoznak a hálózati csomópontoknál, majd a vonalakon megosztva közlekedhetnek tovább. Gyakran előfordul az az eset, hogy egy útvonalválasztó gyorsabban kapja a csomagokat a beérkező vonalairól összességében, minthogy azokat pillanatnyilag ki tudná szolgálni. A kiszolgálás pillanatnyi akadályoztatása mögött két kézenfekvő ok található. Az egyik, hogy a helyes irányvá-

lasztás döntésének meghozásához idő szükséges, a másik kézenfekvő indok, hogy a kimenő vonal pillanatnyilag foglalt lehet. Ilyen esetekben az útvonalválasztó ideiglenes tárolókba, pufferekbe teszi a csomagokat, hogy alkalomadtán később rendre kiszolgálhassa azokat is.



2.1. ábra. Kommunikációs séma egy csomagkapcsolt hálózaton. Két forrás és három nyelő működik egyszerre: $F_1, F_2; Ny_1, Ny_2, Ny_3$. Az $F_1 \rightarrow Ny_3$ és az $F_2 \rightarrow Ny_1$ irányú beszélgetés az X -el jelölt útvonalválasztó pufférében keveredik, csomagjaik egymás mögött sorba kényszerülhetnek. Az $F_2 \rightarrow Ny_2$ irányú beszélgetéshez az A -val jelölt útvonalválasztón több lehetséges út is megvalósulhat.

2.1. A hálózati elemek fizikai paramétereit

Az eddig felmerült hálózati elemeket, fizikai tulajdonságaikat jól megragadó paraméterekkel írhatjuk le. Az összekötő vonalakra jellemző a propagációs idő (jele: τ , mértékegysége: s), amely megadja mennyi idő szükséges ahhoz, hogy a fizikai jel, és így az információ csomagok a vonal egyik végéről a másikra eljussanak. Az útvonalválasztó egységet jól jellemzi a kapacitása (jele: C , mértékegysége: bit/s), az a legnagyobb sebesség, amivel fel tudja dolgozni és kiszolgálja a csomagokat. Az útvonalválasztókba beépített pufferek méretét B -vel jelöljük (mértékegysége: bit , vagy csomag darabszám).

2.2. A kiszolgálási sorokról

A számítógéphálózatok és az Internet valójában egy bonyolult sorbanállási rendszer, amiben a bemenő vonalak, az útvonalválasztó puffere és csomagkiszolgáló egysége, valamint a kimenő vonal együttese egy kiszolgáló egységnek, idegen szóval egy *hopnak* tekinthető. A kiszolgáló egységek viselkedésének teljes leírásához meg kell határozni a beérkezési igényeket, a sort és a kiszolgálást jellemző

paraméterek teljes rendszerét. A sorbanállási rendszerek részletes leírása megtalálható az irodalomban [21], a minket érintő fogalmakat itt tömören vázoljuk.

A beérkezési igényeket legtöbb esetben sztochasztikus felújítási folyamatnak tekintjük, a beérkező csomagok között eltelt t_{cs} idő az $\mathcal{P}(t_{cs} \leq t)$ eloszlást követi, ami felső szinten a felhasználói szokásokból és az alkalmazást jellemző információs igényekből, alsóbb szinten egy-egy hálózati protokoll saját vezérlési tulajdonságaiból együttesen, bonyolult módon áll elő. $\mathcal{P}(t_{cs} \leq t)$ tipikusan Pareto illetve Weibull folyamat.

Abban az esetben, amikor a beérkezési igény nagyobb, mint amit ki tud szolgálni a rendszer, azaz gyorsabban érkeznek a csomagok, mint amilyen nagy a kiszolgálási kapacitás ($C < 1/(t_{cs})_T$), akkor a csomagok sorokba kényszerülnek. Általánosan egy sor kiszolgálása több vezérlési elv szerint is megvalósulhat. Az Interneten egy kiszolgáló egység tipikusan érkezési sorrendben szolgálja ki a sorban álló csomagokat, a *first in first out* elvet követve, de a hálózatban előfordulhatnak olyan sorok, amikben külön megjelölt csomagok nagyobb prioritást élveznek, és a sor elejére kerülnek. Ez utóbbira eklatáns példa a gyors választódott megkövetelő alkalmazások, mint az élő kép- és hangtovábbítás. Az Interneten közlekedő információs csomagok soraira nem jellemző a sorbanállás elméletében sokszor tárgyalt veremszerű kiszolgálás elve (*last in first out* elv), sem a véletlen sorrendű kiszolgálás (*serial in random out* elv).

A várakozási sorok mérete a gyakorlatban véges (B) nagyságú, ha az újonnan érkező csomag nem fér a sor végére, elveszik¹.

2.3. Az egyedi csomagküldési dinamikáról

A hálózatban a kommunikáció jellemzően végpont-párok között valósul meg, az egyik végpont a forrás, a másik a nyelő szerepét tölti be. A forrás végpont feladata felmérni a hálózatban pillanatnyilag rendelkezésre álló erőforrásokat és ennek megfelelően beállítani a csomagküldés sebességét. Ezeket a mérés alapú vezérléseket összefoglalóan a protolloknak hívjuk. Az egyedi csomagküldés dinamikáját a TCP/IP protokoll [22, 23] írja le. A protokoll szerint a csomagküldési sebességet a forrás *terminál* mindaddig növeli, amíg szabad kihasználatlan kapacitást érez a hálózatban és azonnal csökkenti, amint torlódást tapasztal. A TCP/IP protokoll ténylegesen megvalósított algoritmus az évek során sokat finomodott, részletes ismertetése nem képezheti a dolgozat tárgyát. Ebben a részben csak a minket érintő, talán a legelterjedtebb megvalósításban is megtalálható protokoll tulajdonságait vázoljuk fel.

A számítógéphálózatokban minden kommunikálni szándékozó terminál páros először felépíti egymás között a kapcsolatot, majd megindul a tényleges informá-

¹Léteznek olyan sorok (*random early detection* elv), amikben nem feltétlenül az újonnan beérkező csomag veszik el, hanem a várakozási sor végéről, pár csomag mélységig megengedve, véletlenül kilök egy, már eleve ott várakozó csomagot.

ciócsere, mely során a csomagküldés sebességét a TCP algoritmus szabályozza. Kezdetben a forrás csak egyetlen csomagot küld a hálózatba, amit miután a nyelő megkap, egy kis méretű csomaggal (ACK) nyugtáz. Azt az időt, ami egy csomag kiküldése és az arra válaszul küldött nyugtacsomag megérkezése között eltelik, körüljárási időnek (RTT) hívjuk. A szabályzó algoritmus különböző állapotának leírására több belső paraméterrel és változóval is rendelkezik. Egyik fontos belső változója a w ablakméret (*congestion window*), ami egy nyugta beérkezése után a kiküldendő csomagok számát méri. Kezdeti értéke $w = 1$, ami a $w \mapsto w + 1/w$ leképezés alapján módosul, amint a küldő terminál nyugtát kap. Abban az esetben ha az ablakméret w értéke egész számot átlépve változik, két csomagot küld a terminál, ellenkező esetben csak egyet. Így az ablakméret alsó egész értéke ($\lfloor w \rfloor$) számot ad a hálózatba kiküldött de még nem nyugtázott csomagok pillanatnyi darabszámáról. Állandó körüljárási időt feltételezve mind az ablakméret, mind a hálózatban közlekedő csomagok száma időben közel lineárisan nő. Ez a tendencia csak addig tart, amíg valahol, a hálózatban kialakult dugulás következtében el nem veszik egy csomag. A veszteségre reagálva a csomagküldési sebességet csökkenteni kell. A TCP algoritmus a $w \mapsto \beta w$ leképezés szerint, multiplikatív csökkenti az ablakméretet. ($\beta < 1$), és a nyugták érkezésekor mindaddig vár újabb csomagok hálózatba küldésével, amíg a hálózatban kinn levő csomagok száma meg nem egyezik az új ablakméret alsó egész részével. Amint újra teljesül, hogy az ablakméret helyesen méri a hálózatban közlekedő, nyugtázatlan csomagok számát, a csomagküldési sebességet szabályzó algoritmus visszatér eredeti lineárisan növekvő fázisába.

Abban az esetben, ha ablakméret magas értékeket vesz fel, annak lépcsős, granularis természete elhanyagolható lesz, és lehetőség nyílik a változó folytonos közelítésére. A lineáris növekedést leíró állapotban az ablakméretet az jellemzi, hogy minden körüljárási idő elteltével, T_{RTT} időközönként eggyel nő, így az eredeti dinamikát jól közelíti a következő differenciális formula:

$$\frac{dw}{dt} = \frac{1}{T_{RTT}(t)}, \quad (2.1)$$

ahol a $T_{RTT}(t)$ kifejezés a terminál által pillanatnyilag mért körüljárási időt jelenti. Csomagvesztés után az ablakméret a

$$w(t_+) = \beta w(t_-) \quad (2.2)$$

képlet szerint változik. A képletben szereplő $t_{\pm} = t \pm 0$ közvetlenül a csomagvesztés után illetve előtt bekövetkező pillanatot jelenti, valamint $\beta \in [0; 1]$.

2.4. Szimulációs környezetek

A számítógéphálózatok vizsgálatának manapság elengedhetetlen kellékei a szimulációs környezetek. Ezekben valódi hálózati komponensek huzalozása nélkül,

tetszőleges kísérleti hálózat építhető össze, és a hálózaton különböző dinamikájú, az Interneten is működő, forgalomforrás és nyelő működtethető. A legtöbb hálózati szimulátor környezet diszkrét idejű, eseményvezérelt elven működik, azaz az egyedi beépített elemek mozgásegyenleteit numerikus módszerekkel oldja meg, és az összes állapotleíró csak az érdekes időpontokban értékeli ki. Tipikusan érdekes időpont például egy csomag bekerülése egy útvonalválasztó pufferébe vagy egy új alkalmazás bekapcsolása. A szimulátorban megvalósított hálózat a forgalmát leíró paraméterek időbeli fejlődése kényelmesen nyomon követhető és kinyerhető további elemzés céljából. A szimulátorok tervezése mögött alapos mérnöki munka húzódik meg annak érdekében, hogy a valóságos hálózat alkotó elemek tulajdonságait hűen követő, műtermékektől mentes modellt vizsgálhasson a szimulátort felhasználó kutató. A mérnöki irodalomban [24] sokat foglalkoznak a szimulátor hűség kérdésével. Munkánk során mérvadónak fogadjuk el a szimulátor eredményeit, szem előtt tartva annak a lehetőségét, hogy a legfinomabb részleteket tekintve a valóság eltérhet a szimulációs eredményektől. A 6. fejezetben bemutatandó munkánk során a Berkley Egyetemen 1989. óta folyamatosan fejlesztett *ns* hálózati szimulátort használjuk [25].

3. fejezet

A mérőhálózat

3.1. Motiváció

Az élő hálózaton előforduló torlódások folyamatos monitorozó vizsgálata hasznos célkitűzés. A kritikus dugulások észrevétele és helyének meghatározása fontos feladat, mert a torlódások közvetlenül hatással vannak az Interneten működte-tett alkalmazások megbízhatóságára, hatékonyságára. Kutatómunkánk gyakorlati céljául tűztük ki, hogy egyre megbízhatóbb új mérési eljárásokat dolgozzunk ki és egyre pontosabb mérőeszközöket építsünk, melyek együttes használata megteremtik annak lehetőségét, hogy az Internet hálózatban mérhessük a torlódásokat. Remény szerint, mérési eredményeink megbízható és hasznos adatokkal szolgálnak mind a különböző hálózati modellek valós adatokon nyugvó felparaméterezéséhez, mind pedig mint új szolgáltatás pozitív hatással lesznek a jövőben tervezendő alkalmazási protollok vezérlési mechanizmusaira.

3.2. Helyi monitorozó mérések

A méréseket a kísérlet helyének szempontjából alapvetően két osztályba sorolhatjuk, vannak helyi monitorozó illetve vannak végpontmérések.

Helyi monitorozó mérésről beszélünk, amikor a kutató vagy az operátor valamelyik hálózati eszközön futtat mérőprogramokat. A felhasználó rendszerint megfelelő adminisztratív jogokkal bír, és az eszköz állapotát jellemző változókat folyamatosan megfigyelheti, naplózhatja. Így például a kísérletező egy útvonalválasztó eszközön áthaladó forgalom teljes keresztmetszetéből képes mintát venni, befolyással bír a csomagok útvonalválasztását leíró szabályrendszerre, szűrheti az átfolyó csomagforgalmat, nyomon követheti a várakozási sorok viselkedését vagy leolvashatja az eszköz mutatóit. A kísérlet során keletkező mérési naplót helyben tárolja az eszközön, előzetesen feldolgozza, majd az adatok begyűjtése után tovább elemezheti azokat. Fontos megemlíteni, hogy az ilyen helyi kísérletekben keletkező adatok hálózaton keresztül történő rendszeres mozgatása a hálózatban

plusz forgalmi terhelésként jelentkezik, a forgalom eredeti képét torzítja.

A kutatók gyakorlatilag helyi monitorozó mérést soha nem végezhetnek, aminek egyik akadálya az, hogy a hálózat belsejében működő céleszközök erőforrásai, az esetek zömében, nem teszik lehetővé az adatok bonyolult helyi elemzését és a hosszú naplófájlok ideiglenes tárolását. A helyi monitorozó mérések széleskörű elterjedésének másik akadálya, hogy az eszközöket üzemeltető szolgáltató legtöbbször ellenérdekelt abban, hogy a felügyelete alá eső hálózat jóságát feltáró adatok közkinccsé váljanak, valamint a hálózat felhasználóinak jogait szem előtt tartva, eleve nem adhat ki semmilyen forgalmat jellemző információt. Meg kell jegyeznünk azonban, hogy léteznek olyan mérési projektek [26, 27, 28], ahol tudományos indíttatásból ilyen jellegű referencia adatokat publikálnak, például csomagforgalom mintát tesznek közzé. Ilyen esetben a nyersadatokat alaposan anonimizálják [29, 30, 31], vagy megfelelő kumulatív statisztikákat publikálnak csupán, hogy ne lehessen visszaélni a belőle kinyert információval.

3.3. Végpont-végpont mérési módszerek

A hálózati mérések másik osztályát a végpontmérések [32] alkotják. Amint azt az előző fejezetben láttuk, gyakorlatilag a hálózat belsejét leíró paraméterek közvetlen mérésére nincs módunk, mérési eljárásaink segítségével azonban becslést adhatunk ezekről az adatokról is. A következő alfejezetben azokkal a végpont mérési módszerekkel ismerkedünk meg mélyebben, amiket felhasználunk a hálózat belsejében működő sorok dinamizmusának feltárására irányuló munkánk során.

Az ismertetésre kerülő végpontmérések alapegysége a hálózati csomag. A mérések során a mintacsomagok segítségével valamilyen célzott hatást váltunk ki a hálózathoz, amit megfigyelünk, vagy a csomagot érintő különböző késleltetéseket mérjük meg.

Számos hálózati jellemző meghatározása visszavezethető a csomaggal kapcsolatos időmérésekre. A következő időzítéssel kapcsolatos fogalmak merülnek fel egy-egy hálózati csomag élete során. A hálózati késleltetés az az idő, amely ahhoz szükséges, hogy a csomag a feladó termináltól eljusson a cél végpontba. A hálózati késleltetés hordozza magában az összes információt, ami számot ad a csomag és a hálózat valamint a különböző csomagok közti kölcsönhatásokról. A hálózati késleltetést a következő összetevőkre szokás bontani: fizikai késleltetés, sorban állási idő és kiszolgálási idő.

A fizikai késleltetés vagy idegen kifejezéssel *propagációs* idő számot ad arról, hogy egy adott vonalon a fizikai jelek továbbításához mennyi idő szükséges. Hallgatólagosan a propagációs időbe értjük bele a csomagok a vonalhoz kapcsolódó hálózati eszközök áramköreiben elszennvedett állandó időkésleltetését is. Ennek a mennyiségnek a jele dolgozatomban τ .

A sorbanállási idő írja le, hogy mennyi időt várnak az egyes hálózati csomagok az útvonalválasztó elemek puffereiben addig, amíg az előttük beérkezett

csomagokat kiszolgálja az eszköz. Ez az időtartam véletlen változó, számot ad a hálózat terheltségi állapotáról, a dolgozat az 5. fejezetében ezt a mennyiséget vizsgáljuk meg alaposabban, jele x . A sorbanállási idő a hálózat viselkedése szempontjából fontos statisztikus jellemző, számot ad a hálózatban tapasztalható torlódásokról.

A hálózati késleltetés harmadik tagja a kiszolgálási idő, amely megadja, hogy a hálózati eszközben a csomag mennyi idő alatt dolgozható fel, mielőtt a kiemenő vonalon elhagyja azt. A kiszolgálási idő túlnyomó része a hálózati csomag memóriában és a perifériákon történő mozgatásából származik, ezért nagyon jó közelítéssel a csomagmérettel, P -vel arányosnak vehető, azaz $t_s = P/C$, ahol C a kiszolgáló eszköz kapacitása.

A fenti definíciók értelmében a csomag az l szegmensekből felépülő teljes útvonalán, L -en a következő mértékű hálózati késleltetést tapasztalja:

$$\Delta t = \sum_{l \in L} \left(\tau_l + x_l + \frac{P}{C_l} \right). \quad (3.1)$$

3.3.1. Passzív mérési eljárások

A mérési módszereket, a hálózati forgalomra gyakorolt hatása alapján, további két csoportra bonthatjuk: léteznek passzív és aktív mérési eljárások.

A passzív mérési módszerek legfontosabb jellemzője, hogy a mérés nincs semmilyen zavaró hatással a hálózati forgalomra, ilyenkor ugyanis a mérőponton beavatkozás nélkül figyeljük meg a pillanatnyi forgalom különböző tulajdonságait. Lehetőségünk nyílik arra, hogy a mérőpontra érkező, az onnan kiinduló, vagy az azon áthaladó csomagforgalomról teljes, részletekbe menő képet alkothassunk. A legelterjedtebb passzív mérési alkalmazás a `tcpdump` program, aminek segítségével a mérőpont kapcsolatain a teljes csomagforgalmat lehetőségünk nyílik olvasni, felbontva minden csomagot a rendszer órája segítségével. A csomagokat érintő eseményeket μs granularitású időbélyegekkal ruhazza fel a program, de az idő pontossága e felett van, a számítógép és az operációs rendszer megszakításkezelésétől függ az elérhető tényleges pontosság. A nagyságrendek érzékelte végett $1 \mu s$ idő alatt egy átlagos, 100 Mbps kapacitású eszköz ≈ 10 bájtnyi információt szolgál ki, ami körülbelül a legkisebb hálózati csomag negyede. A `tcpdump` program futtatásához a mérőgépen rendszergazda jogokkal kell rendelkezniünk. A program nyersadatainak elemzésével számos érdekes kérdésre kaphatunk választ, melyek közül itt csak pár példát említünk. Milyen a forgalmat alkotó csomagok méreteloszlása? Milyen érkezési statisztikát követnek a megfigyelt csomagok? Melyek azok a tipikus alkalmazások, amik a megfigyelhető forgalmat létrehozzák? A közvetlen kapcsolat rendelkezésre álló sávszélességből mekkora részt használunk fel?

A passzív mérési eljárások kétségtelen előnye, hogy nyersadatai és a belőle származtatott eredmények mentesek a mérés zajától, ám csak egyetlen hálózati

pontra vonatkoznak, és így általánosabb kérdéseket nehezen vagy csak több mérőpont együttes alkalmazásával lehet megválaszolni velük. További hátránya a módszernek, ahogy a 3.2. részben bemutattuk, hogy a hálózat belsejében üzemeltetett eszközökhöz nincs az átlagos kutatónak hozzáférése, mert például a mérési nyersadatok nem körültekintő feldolgozása visszaélésekre adhat lehetőséget, a felhasználók személyes jogait sérti, valamint a szolgáltató esetlegesen kockáztatja a biztonságos üzemeltetés feltételeit.

3.3.2. Aktív mérési eljárások

Aktív mérési eljárásnak nevezzük azokat a módszereket [33, 34], amikor a mérés során ismert, általunk mesterségesen előállított próbaforgalommal zavarjuk meg a rendszer működését, és megvizsgáljuk a hálózati elemek válaszát a beállított próbaforgalomra. E tekintetben az aktív kísérletek a fizikai szórás-problémákra hasonlítanak a legjobban.

Az aktív módszer egyik fontos előnye, hogy csak a mérésben érintett végpontokhoz szükséges hozzáféréssel rendelkezünk. A próbaforgalom felépítésénél gyakorlatilag minden paramétert könnyen változtathatunk, bonyolult csomagmintázatokat építhetünk, melyek mérési céljainknak leginkább megfelelnek. Az aktív mérések könnyen megismételhetők, ugyanazzal a próbaforgalommal lehetőség nyílik a hálózat különböző időbeli állapotának vizsgálatára.

A próbaforgalom tervezésénél és felépítésénél azonban nagyon körültekintően kell eljárni. Az aktív mérésekre jellemző, hogy a mérőforgalom, a hálózat működése szempontjából, nem különbözik a kommunikációs folyamatok csomagforgalmától, azzal kölcsönhat, s a mérés kibillentheti a hálózatot a mérés előtti állapotból, mérési műterméket idézve elő. A konkrét próbaforgalom sűrűségének megválasztásánál meg kell találni annak optimális értékét: egyfelől a próbák között hagyni kell arra időt, hogy a rendszer relaxáljon, másfelől túl ritka minta esetén a mérés időben elnyúlik, ami akkor jelent problémát, ha a vizsgált tulajdonság állandóságának időskálája kicsi a mérés időtartamához képest, azaz a rendszer vizsgálni kívánt állapota, a mérés során megváltozik. Erre sarkított példa, ha túl ritkán mérünk két pont között, nem garantálható, hogy a mérőcsomagok ugyanazt az utat járják be. A próbacsomagok felépítésénél és küldésénél szintén fontos szempont, hogy a próbák ne tűnjenek támadó forgalomnak, ellenkező esetben a mérésben részt vevő végpont forgalmát a szolgáltatók korlátozhatják, rosszabb esetben teljesen leválaszthatják a hálózatról.

Érdeemes megemlítenünk, hogy az Európai Akadémiai Hálózat gerinchálózatan üzemelő útvonalválasztókon weben keresztül elérhető számos aktív mérési eljárás [35], melyek segítségével egyszerű mérési feladatokat kezdeményezhetünk úgy, hogy a kiválasztott router a mérés egyik végpontja.

A hálózat torlódási mintázatát feltáró aktív mérési alkalmazásokat a következő szakaszban vizsgáljuk meg kicsit részletesebben.

3.3.2.1. Egy útvonal feltárása a hálózatban

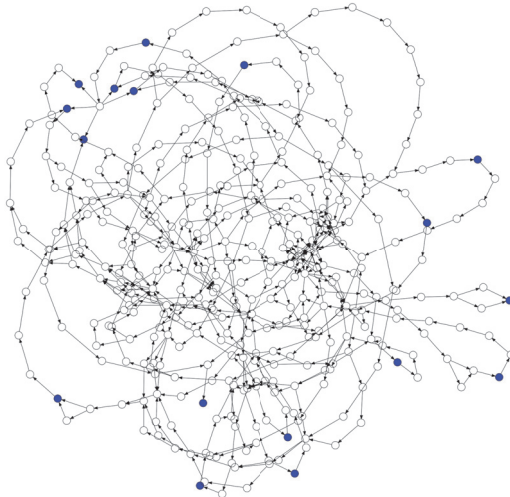
Amint azt korábban említettük, a megfelelő címmel ellátott csomag hálózati bolyongása során az útvonalválasztó egységekben helyi döntések alapján kerül kiválasztásra az, hogy a csomag melyik úton haladjon tovább, ha több kiemenő vonal is lehetséges. Legelterjedtebben a legrövidebb utak elve teljesül, de vannak kivételek, illetve vannak esetek, amikor több legrövidebb út is kijelölhető két végpont között. Sokszor érdekel minket az útvonal teljes belső szerkezete, nem elégszünk meg azzal, hogy ismerjük végpontok között megvalósuló utak hosszát. Az utak mikroszkopikus szerkezetének ismerete elengedhetetlen, például ha több útvonalat összekomponálva meg akarjuk vizsgálni a hálózat – a mérés által feltárt – részének gráfját, annak geometriai jellemzőivel, úgy mint a feszítő hálózatot, a fokszámeloszlását, a betweeness centrality-t.

A legelterjedtebben használt útvonal feltáró alkalmazás a **traceroute** program. A mérés elvégzéséhez elegendő egy végpont. A program segítségével a végpontból kiinduló utakat tudjuk mélységében feltárni. Az útvonal feltárása iteratív eljárás keretében zajlik, a kiszemelt végpont irányába egyre mélyebben nézünk a hálózatba. A hálózati csomagok fejlécében a címeken felül és sok egyéb más fontos paraméter mellett, megtalálható az élettartamot leíró mező¹, aminek értékét a csomag útja során az útvonalválasztó eszközök egyesével csökkentik. Amikor az élettartam mező értéke nullává válik, az útvonalválasztó a csomagot eldobja, és erről értesíti a küldő alkalmazást. Az értesítésbe becsomagolja a saját címét, és ezt az információt használjuk fel az útszakaszok szélső pontjainak azonosítására. A **traceroute** program az élettartam mezőt minden mérési körben növeszti, és az eldobott csomagok után érkező értesítések alapján előállít egy címsorozatot, amiből felépíthető az útvonal. A program ennél picivel több nyersadattal szolgál, szerény *ms* nagyságrendű időfelbontással számot ad a körüljárási időről is. Az alkalmazást több hálózati protokollra is megvalósították.

A 3.4. szakaszban bemutatandó mérőhálózatunk rendelkezésre álló végpontjai között rendszeresen feltárjuk a köztük pillanatnyilag megvalósuló hálózati összeköttöttséget. A 2007. január 11-i, kiragadott példamérésben 12 forrás végpont vett aktívan részt. A források a **traceroute** program segítségével tárták fel az addig telepített 17 végpont felé az útvonalakat. Az útszakaszokat helyesen egymás után fűzve feltártuk a végpontok között feszülő topológiát, melyek egy irányított gráfot alkotnak. A 3.1. ábrán szemléltetjük az így feltárt hálózat komplexitását. Az ábrán kék korongok jelölik a végpontokat, az élek irányítását nyilak jelzik. A hálózatot 409 csúcs között futó 562 egyedi él alkotja, melyek nagy része az Európai Akadémiai Hálózat nagy sebességű vonalainak felelnek meg.

A feltárt hálózatra megvizsgáltuk az utak hosszának eloszlását is. A 3.2. ábra szerint azt tapasztaltuk, hogy a feltárt hálózat a több száz építő él és csúcs ellenére, a kisvilág hálózatokra jellemző kis átmérővel bír. Az utak hossza jó közelítéssel követi a normális eloszlást, a 14,7 várható érték jó összhangban van az

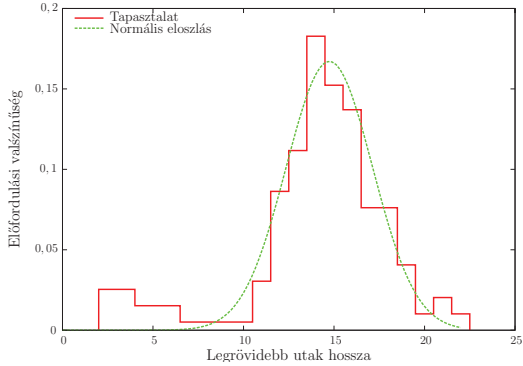
¹ *TTL* – time to live



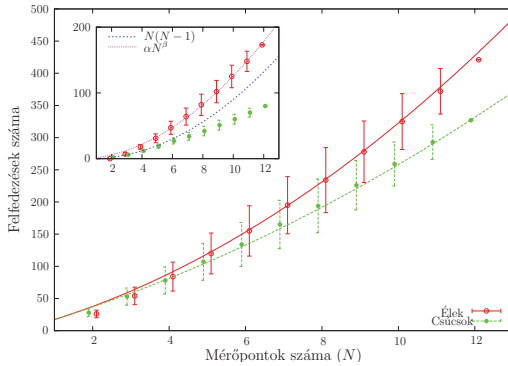
3.1. ábra. A 2007. január 11-én 17 Etomic végpont között leveleznyelt **traceroute** mérésből előállított topológia.

Internetre nagyobb részét felölő topológiai elemzésekkel. Figyelmes szem egy kisebb csúcsot fedezhet fel a 4 hosszúság környékén, ami mérési hibának tekinthető. Léteznek olyan útvonalválasztók, amelyek nem kooperatívak, eltítkolják címüket, vagy szűrik az úton náluk távolabbi útvonalválasztók válasz üzeneteit. A zsargonban csillagosnak nevezik az olyan útvonalat, ahol annak csak a töredéke ismerhető meg, vagy benne ponthibák fordulnak elő. Ezeket a hibákat pontos topológiai elemzéseknel figyelembe kell venni, a nyersadatokat korrigálni vagy szűni kell.

Érdeemes megvizsgálni, hogyan skáláz a hálózatot alkotó csúcsok illetve élek száma, ha növeljük a mérésben résztvevő végpontok számát. A 3.3. ábrán ezt a jelenséget vizsgáljuk. Az elemzés során megfelelően szűkítjük le a mérési nyersadatokat. A forrás és végpontokat magába foglaló 12 elemű metszethalmazból húzunk véletlen $n \leq 12$ elemet. A mérésből ezután csak azokat az utakat hagyjuk meg, melyek kezdő vagy végpontjai a sorsolt elemek közül vannak. Az utakból felépítjük az n végpont között feltárható hálózatot, melynek számoljuk csúcsait és éleit, majd n különböző realizációjára sokaság átlagot képzünk. Az ábra beékelt részén a felfedezett topológia feszítőfákra lesűkített méretfüggését mutatjuk be. Az ábrán feltüntetjük a teljes összekötöttséggel rendelkező gráf skálázásának háttérét. Ebben az esetben az élek száma a rendszer méretével $\approx N^2$ nő. Jól



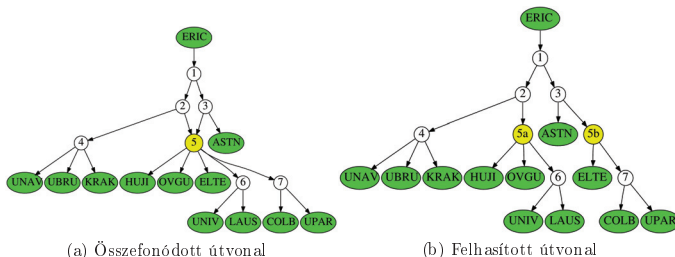
3.2. ábra. A 2007. január 11-én végzett **traceroute** mérések elemzése. A végpontok között feszülő utak hosszának eloszlása, melyet Gauss-eloszlással közelítve az úthossz várható értéke $m = 14,7$, ami jellemző a teljes Internetre is. A szórás kísérleteink alapján $\sigma = 2,4$ -nek adódott.



3.3. ábra. A 2007. január 11-én végzett **traceroute** mérések elemzése. A feltárható topológia csúcs- és élszámainak növekedése újabb végpontok hozzáadásával. A felfedezhető élek számát αN^β hatvány-függvénnyel illesztjük, a paraméterekre $\alpha = 1,55$ és $\beta = 1,91$ értékeket kapunk kisebb, mint 10 %-os hibával.

látszik, hogy a felbontható szegmensek száma szignifikánsan nagyobb, mint a végpontok között megvalósuló utak száma. Hatvány-függvény szerinti viselkedést feltételezve $\beta = 1,9$ -es kitevőt illesztünk, melyből a végpontok duplázásával várhatóan az belső élek száma 3,8 szerezére nő.

Topológia feltáró méréseink során, 2006 előtt az eredeti **traceroute** programot használtuk. Ez időtájt azonban az Internet általunk vizsgált részén egyre több olyan új eszközt építettek a hálózatba, mely lehetőséget teremt az azelőtt nem tapasztalt dinamikusabb útkiosztásra, mely mögött redundáns vonalkiépítés húzódik. Azt tapasztaltuk, hogy a méréseket rendre egymás után megismételve, más-más útvonalat kapunk eredményül. Az 5. fejezet módszerei azonban állandó és ismert hálózati összekötöttséget követelt meg, ezért az utak feltérképezésére más megoldást kellett keresnünk. Felismertük, hogy ha rögzítjük a csomag címzésének finom részleteit (UDP protokollra tértünk át és rögzítettük a címzésben szereplő portpárt is), a módosított programmal továbbra is megbízható marad az útvonalfeltárás, stabil reprodukálható utakat mérünk. Kísérletező kutatásainknak ezen a pontján azonban elkövettünk egy hibát. Megelégedtünk azzal a ténnyel, hogy a torlódás vizsgálatára bevett módszereinkhez továbbra is elő tudjuk állítani a szükséges hálózati összekötöttséget, s csak kis hangsúllyal foglalkoztunk az alternatív utak kérdésével. Megszerzett tapasztalatainkkal nem tudtunk érdemben hozzájárulni e jelenség feltérképezésére kifejlesztett mérőprogram, a **paristraceroute** megalkotásához [36].



3.4. ábra. Példa az összefonódó utakra. Az 5 jelölt útvonalválasztónál újra összefolynak az utak, amiket figyelembe véve, hogy melyik levélhez pontosan merre vezetett az eredeti út, $5a$ és $5b$ klónok bevezetésével és az utak helyes kiosztásával megfelelő feszítőfa szerkeszthető.

Az 5.3. rész fás tomográfiai módszeréhez a feltárt útvonalakból feszítőfákat kell építenünk. Minden feszítőfa egy gyökér végpontból indul, ami megfelel a mérőforgalom forrásának, és levelekben végződik, amik szintén a mérésben résztvevő gépek. A fát alkotó belső pontok elágazási pontok, az élek összetettek: az eredeti

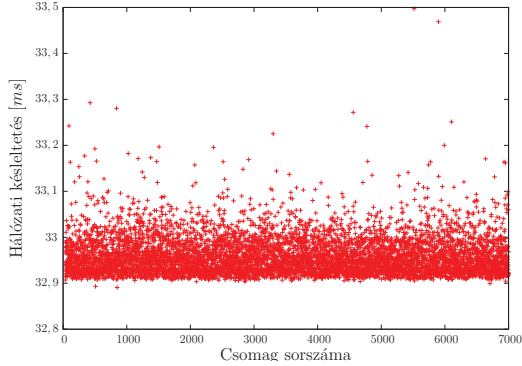
szakaszok olyan láncolatai, melyben a szomszédos egy-fokú csúcsok vannak fel-fűzve. Az alternatív utak problémája miatt azonban a feszítőfa megszerkesztése nem csupán az egy-fokú csúcsok kiredukálása az eredeti hálózataból. Ritkán előfordulnak olyan esetek, amikor az élek újra összefonódnak, lásd a 3.4. ábrán. Ilyen esetben a megfelelő csúcsok klónozásával tudjuk fává alakítani a feszítőhálózatot. Az összefonódó utaknál mindig az első elágazási pontot tekintjük az *igazi* elágazásnak, mert a mérőcsomagok ezen a ponton dekorrelálódnak.

3.3.2.2. A hálózati késleltetés mérése

Az előző pontban bemutatott méréseknél, az útvonalak feltárásánál a hálózat belsejében működő elemek kooperálására van szükség, hiszen a mérés során hálózatra küldött próbák hatására új csomagok keletkeztek az útvonalváltásokban, melyeket a végponton elkapva feltárjuk, merre haladt nagy valószínűséggel a próba.

Számos esetben a csomagok időzítéseit szeretnénk megvizsgálni, leginkább arra vagyunk kíváncsiak, hogy külön-külön mekkora az egyes hálózati elemeken elszennvedett késleltetés, amire általában nincs közvetlen lehetőségünk. Általánosan méréseink csak egy teljes útvonalon tapasztalható teljes késleltetésre irányulhatnak, hiszen legtöbbször csak azt mérhetjük meg, hogy a próba mikor került a hálózatba illetve onnan mikor érkezett meg a célponthoz. A próbacsomagok mintázata alapján megkülönböztetünk egy- és több-csomag méréseket. A legegyszerűbb egy-csomag késleltetés mérő alkalmazás a *ping* program, amivel a körüljárási idő határozható meg. A körüljárási idő a 3.1. képlet alapján gyakorlatilag úgy áll össze, hogy az utat alkotó élek L halmazba beleértjük a válaszcsoomag útvonalát is, és a kifejezést kiegészítjük egy új véletlen változóval, ami megadja, hogy az megcélzott végpontnak mennyi időbe telik elkészíteni a válaszüzenetet. Legtöbb esetben feltételezzük, hogy az útvonal szimmetrikus és elhanyagolható a csomag elkészítési ideje, ezekben a modellekben a körüljárási időt megfeleltetik a hálózati késleltetés kétszeresének.

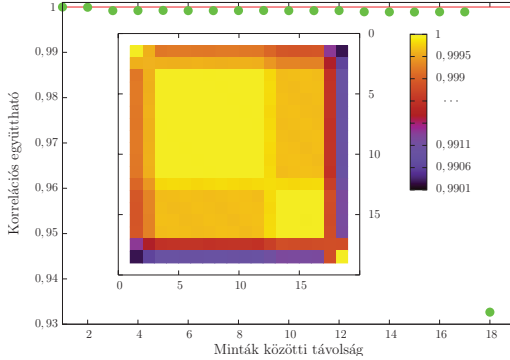
A több-csomag mérésekben hálózatra küldött minta, az egyedi idősorok meghatározásán felül, egzotikusabb adatok megmérésére is lehetőséget nyújt. Miközben áthaladnak a hálózaton, a mintában elhelyezett csomagok egymáshoz viszonyított helyzete is változik, amiből érdekes következtetéseket lehet levonni. Például előre meghatározott időközlelt útjára indított csomagpárok vagy csomagtrillák segítségével becslést adhatunk az útvonalon elérhető szabad sávszélesség nagyságára [37]. Azok az eljárások, amik a mintát alkotó csomagok relatív távolságadataira épülnek nem követelik meg a végpontok szinkronizációját, és az órák csúszásával szemben is toleránsak, hiszen a hétköznapi számítógépekbe épített órák, egy minta kiküldési ideje alatt nem csúsznak el számottevően. Szemléltetésül a csomagmérettel összemérhető, $100\ \mu\text{s}$ nagyságrendű csomagszeparációval felépített mintában, átlagosan pontatlan, $20\ \mu\text{s/s}$ csúszású óra alapján elhanyagolható a mérési hiba, csak a harmadik tizedes jegyben történik változás.



3.5. ábra. Hálózatkésleltetési idősor egy 2007. február 18-i mérés alapján. A *jerusalem.etomic.org* mérőpont 11 másik rendelkezésre álló végpontra küld 0,1 s rátával sűrűn egymás mögé helyezett mintacsomagokat. A csomagok indulási és érkezési idejeit pontosan mérjük, és hálózati késleltetés idősorokat állítunk elő. Az idősorok közül egy példa, ami az *elte.etomic.org* végpontra érkezett.

A dolgozatban az 5. fejezetében megvizsgált hálózati torlódásokról számot adó, sorban állási idő tomográfias módszerekhez elengedhetetlen a korrelált csomagmintázat késleltetéseinek precíz mérése. A feladat megoldására saját csomagküldő és -fogadó programokat fejlesztettünk ki, amely a következő 3.4. alfejezetben ismertetett Etomic mérőhálózat végpontjain futtathatóak. A mérőprogram szorosan egymás mögé sorakoztatott (*back-to-back*) csomagokat küld a hálózatba. A mintát alkotó csomagok címzettjei különböző végpontok, így gyakorlatilag a próbacsomagok egy fát járnak be, aminek a gyökere a küldő végpont, a levelei pedig a fogadó mérőpontok. Ilyen fára láttunk példát a 3.3. ábrán. A mérési nyersadatot csatolt idősorok alkotják, amire szemléltetésül a jeruzsálemi és az egyik budapesti mérőpont között mért hálózati késleltetésből a 3.5. ábrán mutatunk részletet.

A csomagküldő program elkészítése során arra törekedtünk, hogy minél kevesebb számolással lehessen a csomagokat előállítani, és a szükséges ellenőrző összegeket előre kiszámoljuk. A kész mintázat küldése a lehető legnagyobb prioritással, a fizikai réteghez legközelebbi szinten történik. A pontos megbízható küldés alapvetően fontos, mert a tomográfias becslési módszerek feltételezik, hogy egy mérőmintázaton belül koherensen küldjük a próbacsomagokat. A küldő alkalmazás jószágát kalibrációs mérésekkel ellenőrizzük. A kalibráció során 20 csomagból felépített mintákat küldtünk a hálózatba és a csomagok indulási időbélyegeit feljegyeztük. Az i -ik minta $j \in [1, 20]$ csomagjának indulási idejét jelölje $t_j^{(i)}$. Elkészít-



3.6. ábra. A csomagküldő program jósági tesztje. A csomagküldő alkalmazástól elvárjuk, hogy az egy mintába tartozó próbacsomagokat a lehetséges legkisebb közbeiekelt időközszel küldje a hálózatba. Egy 20 csomagból álló minta küldési időbélyegei között definiált különbségek korrelációs együtthatóit értékeljük ki. Ideális esetben 1 értékű együtthatót várunk, amit a csomagok közé beekelődő, véletlen idő ront. A külső ábra alapján azt tapasztaljuk, hogy a minta első két csomagja illetve az első és $j + 1$ -ik csomag közötti küldési idők jól korreláltak, a 18. csomag után minőségi romlás tapasztalható, de még így is jól korrelált a teljes csomagvonal. A belső ábrán a korrelációs együtthatók mátrixát is ábrázoljuk.

jük a csomagok indulási idejei között kimérhető időkülönbséget, $\delta_j^{(i)} := t_{j+1}^{(i)} - t_1^{(i)}$, ami ideális esetben állandó, jP/C nagyságú. Az idősorokra kiszámolt korrelációs együtthatók ϵ mátrixát ² a 3.6. ábrán tüntetjük fel. Az együtthatók értéke közel esik a tökéletesen korrelált küldésnek megfelelő 1-hez, ami alátámasztja, hogy a küldőprogramunk ellátja feladatát, a későbbiekben élhetünk azzal a feltételezéssel, hogy a szoros mintában elhelyezett csomagok egymás után elegendően nagy koherenciával kerülnek a hálózatba.

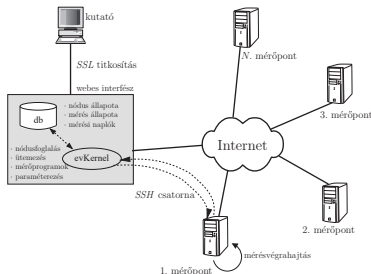
3.4. Az Etomic mérőhálózat

Az Európai Unió VI. keretprogramjában³ meghirdetett Evergrow [38] alprogram részeként, nemzetközi kollaborációban megvalósítottuk a European Traffic Observatory Measurement InfrastruCture [39], röviden *Etomic* mérőhálózatot. A mérőhálózat célja, hogy nagy időzítési pontossággal bíró, az Internetbe ágya-

² $\epsilon_{j,k} := \text{Cov}(\hat{\epsilon}_j, \hat{\epsilon}_k) / \sigma[\hat{\epsilon}_j] \sigma[\hat{\epsilon}_k]$

³IST FET Complexity

zott kísérleti eszközt nyújtson. A mérőhálózat felhasználója lehet minden olyan kutató, aki megfelelő kutatási programmal pályáz a rendszer használatára, és elfogadja annak működési feltételeit.

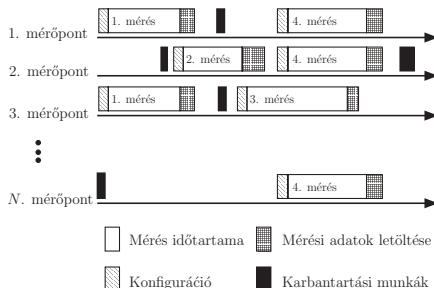


3.7. ábra. A Központi Vezérlő Rendszer (CMS). Az európai akadémiai hálózat peremén elhelyezett Etomic mérőpontok időben szinkronizált hálózati mérésekre alkalmasak. Irányításuk a CMS feladata. A Központi Vezérlő Rendszer felel a mérések ütemezéséért valamint a mérési adatok begyűjtéséért, melyet a kísérletező felhasználó WEB-es felületen ér el további analízis céljából.

Az *Etomic* [40, 41] rendszer két fő összetevőből áll. A Központi Vezérlő Rendszer⁴, (3.7. ábra) feladatában jól elkülönül a mérőpontoktól, többek között ez utóbbiak vezérléséért felelős. A rendszer tervezésekor fontos szempont volt, hogy a hálózati mérésekben jártas és kevésbé jártas kutatók egyaránt megtalálják azokat a funkcionálisokat, amelyek méréseken alapuló kutatásaikat segítik. A Központi Vezérlő Rendszer a felhasználók számára a web-es portálon keresztül érhető el, így különleges programok telepítése nélkül egy mindennap használt böngészőből könnyen konfigurálható minden mérőpont. A portál számos beépített mérőprogramot felkínál használatra, például a *ping* alkalmazás vagy a *traceroute* program több variánsa is elérhető. A rugalmasság érdekében a kutató által kifejlesztett tetszőleges mérőprogram felvezethető a rendszerbe és futtatható, akár úgy is, hogy a mérőprogramnak csak a forrását tölti fel a felhasználó, és abból a mérés megkezdése előtt a rendszer futtatható állományt fordít. A Központi Vezérlő Rendszer segítségével időzíthetjük mérési feladatainkat, majd amikor azok befejeződnek, a mérési naplót a rendszer összegyűjti és felületet biztosít az adatok utólagos letöltésére. Fontos megemlíteni, hogy a megbízható nagy precizitás érdekében azt a filozófiát követjük, hogy egy adott mérőpontot egyszerre csak egy felhasználó használhat. Ellenkező esetben nem garantálhatnánk, hogy különböző alkalmazások egyszerre férjenek hozzá a mérőpont perifériáihoz így leginkább a

⁴ CMS – Central Management System

hálózati kártyákhoz. Ebből következik az is, hogy a Központi vezérlő rendszer felel azért, hogy a mérőpontok helyes ütemezés szerint működjenek, térben és időben különböző mérések ne fedjék át egymást, lásd a 3.8. ábrán.



3.8. ábra. A Központi Vezérlő Rendszer (CMS) mérésütemezési vázlata. A CMS titkosított csatornákon keresztül hozzáfér minden mérőponthoz. A mérésütemezés felelős azért, hogy különböző mérések ne fedhessenek át. Minden mérés megkezdése előtt feltölti a megfelelő mérőprogramokat és azok paraméterezését ütemezését elvégzi, majd mérés végeztével begyűjti az adatokat.

A Központi Vezérlő Rendszer a mérések megkezdése előtt titkosított csatornákon keresztül minden, a mérésben érintett mérőpontra feltölti a megfelelő mérőprogramokat és azok paraméterezését, ütemezését elvégzi. A vezérlőrendszer a mérés ideje alatt nincs kapcsolatban a mérőgéppel, ezzel kívánjuk lecsökkenteni a kapcsolattartás csomagforgalmából eredő, a mérést is befolyásoló zajt. A mérések befejezése után a Központi Vezérlő Rendszer begyűjti a mérési adatokat. A letöltés folyamata alacsony prioritású, ha a letöltés ideje alatt mérést regisztrál a rendszer, akkor a letöltés folyamata ideiglenesen megszakad, majd a mérőgép üresjárás idejében folytatódik minden függőben levő fájlcsere.

A felhasználói portálon keresztül ellenőrizhető a mérőpontok foglaltságán felül azok egészségi állapota is.

A rendszer másik izgalmas elemét maguk a mérőpontok képezik, amik továbbfejlesztett szabványos kiszolgáló PC-k. A mérőgépet speciális hálózati mérőkártyával egészítettük ki a precíz csomagküldés és csomagfogadás érdekében. A mérőkártyát az Új Zélandon működő Endace [42] gyártotta. A cég profilja olyan mérőkártyák tervezése és építése, ami nagy megbízhatósággal, pontos időbélyegzéssel képes hálózati vonalakat vagy *switcheket* folyamatosan monitorozni, akár a teljes csomagforgalom elmentésével. A kártyával úgy érhető el a rendkívül pontos időbélyegzés, hogy a rá épített elektronika a befoglaló számítógéptől nem kér megszakításokat, a csomagokat az első bit beérkezésekor saját órája alapján bélyegzi

fel, saját memóriájában tárolja, és a rendszerbuszon keresztül közvetlen memória hozzáféréssel⁵ írja a PC merevlemezére. Az Endace készít olyan kártyákat is, amik elmentett csomagtraceket képesek visszajátszani a hálózatba. Kifejezetten a projekt részére új kártyatípust specifikáltunk, amit *DAG3.6GE* fantáziánéven gyártottak le a mérőgépekre szabva. A kártyával lehetőség nyílik hálózati csomagok elkapására és azok pontos időbélyegzésére, valamint használatával lehetőség van egyszerű csomagmintázat generálására, identikus mérőcsomagok precíz időközű kiküldésére. Laboratóriumi körülmények között végzett vizsgálataink alátámasztják, hogy az időbélyegzésre használt óra granularitása 100 ns alatt van. A mérőgép időzítési tulajdonságait feltáró mérésben üres vonallal kötöttünk szembe két kártyát, és előre beprogramozott időközönként küldtünk mintákat. Az állandónak várt propagációs idő kimért ingadozásából megállapítottuk, hogy az időbélyegek hibája $\approx 0,5\text{ }\mu\text{s}$. A speciális hálózati kártyával felépített mérőszámítógépek belső referencia órajelét és a földrajzilag távol eső pontok óráinak szinkronizációját gps vevő segítségével oldottuk meg. A kategóriájában viszonylag olcsó Garmin *35HSV* típusra esett választásunk, melynek nagy stabilitású 1 Hz -es négyszögjelét⁶ használjuk fel mind a mérőgép saját órájának, mind a *DAG* mérőkártyán található óra finomhangolására. A gps antenna és a mérőszámítógépek közötti távolságot áthidalandó, a gps eredeti, *RS232* szabványú jelszintjeit saját tervezésű konverter elektronikák segítségével átalakítjuk *RS422* szabványúra. Így több száz méter kábeltávolságra üzemeltethető be a mérőgép a gps antennától. A számítógépeken *Debian linux* operációs rendszer fut, a mérőalkalmazásokat linuxos környezetben kell kifejleszteni a kutatóknak. A *DAG* kártyához fejlesztetni kívánt



3.9. ábra. Az Etomic rendszerben üzemelő mérőpontok földrajzi elhelyezkedése Európa szerte. A mérőgépek listája: *{elte, colbud, univet, pamplona, birmingham, magdeburg, jerusalem, rome, sics, ericsson, telia, krakow, paris, brussels, lausanne, chania, salzburg, skopje}*. etomic.org

⁵*DMA* – Direct memory access

⁶*PPS* – Pulse per second

alkalmazásokhoz *C* nyelvű programozási felületet biztosítunk (API).

A mérőgépeket a 3.9. ábrán feltüntetett különböző európai egyetemen vagy kutatási központban telepítettük. A nódusok jellemzően a gyors elérési hálózat⁷ peremén üzemelnek.

3.5. A Dimes program és a PlanetLab hálózat

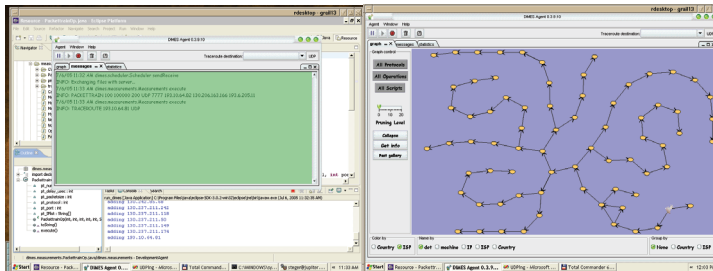
Az 5. fejezetben részletezett, a hálózati torlódásokról számot adó, sorbanállási idő tomográfias méréseinket rendszeresen az EtoMic mérőhálózaton végezzük. Az EtoMic végpontok között az Interneten feszülő hálózatnak csak kisebb kiterjedésű – körülbelül 180 szegmenst magába foglaló – tartományát tárhatjuk fel, ezért fontosnak tartottuk megvizsgálni, milyen már meglévő lehetőséggel terjeszthetjük ki méréseinket és becsléseinket a hálózat olyan részeire is, amit nem érhetünk el a rögzített helyre telepített EtoMic végpontokból.

A méréseink szélesebb területre való kiterjesztésére két lehetőséget találtunk és vizsgáltunk meg. Az első a PlanetLab [43] infrastruktúra, melyet 2003-ban azért hívtak életre, hogy új, elosztott alkalmazások⁸ teljesítképességét globális méretekben lehessen az infrastruktúra segítségével tesztelni. A rendszer méretei és a földrajzi lefedettség tekintve vonzó volt kísérleteink számára: jelenleg 485 intézményben majdnem 1000 végpontot üzemeltet. A PlanetLab működési elve erősen eltér az EtoMic rendszerétől. A végpontokon virtuális gépek üzemelnek. Minden felhasználó egy teljes környezetet kap, ami felett rendszergazda jogokkal rendelkezik, s így elviekben precíz időzítésű csomagmintázat küldése is joga van. Azonban az egymás mellett párhuzamosan működő virtuális rendszerek a valódi erőforrásokon és perifériákon időszeltekben osztoznak. A felhasználónak nincs lehetősége abban a pillanatban időbélyeget rendelni egy próbacsomaghoz, amikor az a végpontot valójában elhagyja, hanem időbélyegei arról adnak csak számot, hogy mikor kerülnek csomagjai egy a virtuális gép és a hálózati csatoló közé beiktatott sorba. A közbülső sort jellemző megbízható időzítési információkkal nem rendelkezünk, e sorok dinamikája erősen függ attól, hogy mennyi virtuális gépet kell egy Planetlab végpontnak kiszolgáltatnia és egy pillanatban hány felhasználó osztozik a hálókártyán. Kutatásunk e szála negatív eredménnyel zárult a tomográfias mérések szemszögéből nézve. Azonban megismerve ezt a rendszert más, a dolgozatban nem részletezett, kutatásba kezdtünk, például szélesebb körben vizsgálhatóvá vált a hálózati topológia geometriája.

Mérésbővítési erőfeszítéseink másik szála a Dimes [44] mérőprogram vizsgálata esett. A Dimes hálózati mérőprogram kifejlesztéséért felelős pályázat szintén az Európai Unió VI. keretprogramjában futott. Alkalmunk nyílt aktív munkakapcsolatot kiépíteni a Tel-aviv Műszaki Egyetemen és a Jeruzsálemi Egyetemen dolgozó fejlesztő kutatókkal.

⁷access network

⁸peer-to-peer alkalmazások



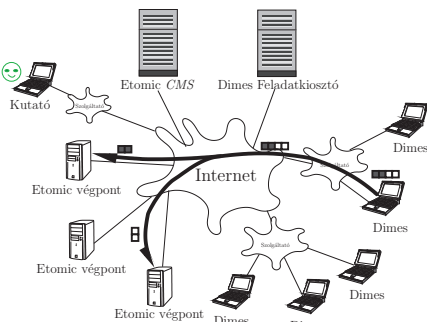
(a) A mérőmodul tesztelése

(b) A referencia mérés topológiája

3.10. ábra. A Dimes mérőprogram. (a) A mérőprogramba saját fejlesztésű csomagmintaátküldő modult illesztettünk, az ábrán ennek tesztelése látszik. (b) A mérőprogram grafikus felülete megjeleníti az Etomic végpontok irányában látszó, **traceroute** mérésekből kirajzolódó topológiát.

A Dimes mérőprogramot eredetileg a hálózati topológia felkutatására fejlesztették ki. Az alapötlet a következő volt: a hálózati összeköttetés felderítéséhez nem kell feltétlenül drága végpontokba befektetni, hiszen úgy csak kevés pontról körbetekintve lehetséges a hálózatot feltérképezni. Célravezetőbb, ha nagy számú olcsó mérőpontot szórunk szét a hálózaton, és még jobb ha ezek a mérőalkalmazások nincsenek helyhez kötve. A Dimes projektben a drága infrastruktúra helyett inkább egy jól reklámozott, mindenki által elérhető alkalmazást fejlesztettek ki. Gyakorlatban a *C* és *java* programozási nyelveken megírt alkalmazás a méréseket a felhasználó számítógépén alacsony prioritással a háttérben futtatja, ezzel kevés erőforrást használ fel azon, valamint törekszik a lehető legkevesebb hálózati forgalom előállítására. A mérőprogramban eredetileg a már ismertetett **traceroute**-ot és **ping**-et implementáltak, melyek segítségével a környező hálózati elemekről nyerhető ismeret. A program grafikus felhasználói felülettel rendelkezik, amin nyomon követhető, milyennek látható a hálózat a felhasználó számítógépének szemszögéből, lásd a 3.10. ábráin. A működés menetrendje szerint minden mérőprogram indulásakor kapcsolatba lép a Dimes Feladat kiosztó Központtal, és egy címlistát kap. A lista elemeit rendre megvizsgálja, körúljárási idő statisztikát gyűjt, és felfedezi a kapott cím felé vezető úton fellelhető útvonalválasztók azonosítóit. Az alkalmazás visszaküldi a Feladat kiosztónak a mérési nyersadatokat, ahol azok későbbi elemzés céljából adatbázisokba kerülnek és archiválódnak.

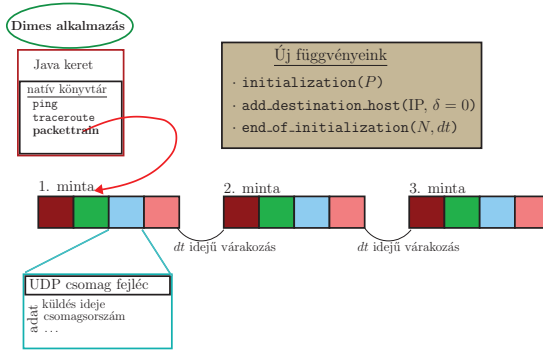
A Dimes program fejlesztői örömmel fogadták ötletünket, hogy kiegészítsük alkalmazásukat új mérési algoritmusokkal, aminek sematikus vázlata a 3.11. ábrán látható. Megkezdődött a közös fejlesztőmunka időszaka. Elsőként olyan modult



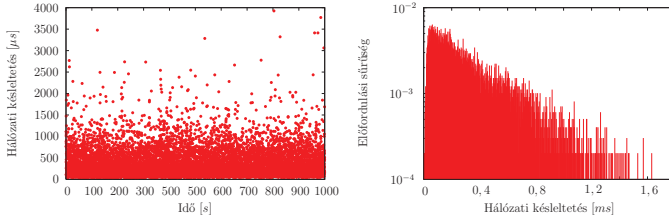
3.11. ábra. A Dimes–Etomic tomográfias mérési elrendezés vázlata. A kutató számítógépéről vezérli a mérést, az Etomic Központi Vezérlő Rendszerben (CMS) kiválasztja a mérés végpontjait és azokat hallgató üzem módba kapcsolja, egyidejűleg a Dimes Feladat kiosztón keresztül egy újonnan bekapcsolt mérőprogramnak tomográfias mintázatküldő programot ír elő. A kutató a mérés végén a CMS-en keresztül éri el a mérés során begyűjtött idősorokat, míg a topológiaadatok a Dimes archívumából kérhetők le.

terveztünk (lásd 3.12. ábra), melynek segítségével lehetőség nyílik arra, hogy a mérőprogram a hálózatba szorosan egymás után, különböző hálózati címmel el látott, csomagmintákat küldjön.

A tomográfias kiértékelés egyik bemenetét a hálózati késleltetés idősorok képezik, és ez rögtön felvet egy problémát. A Dimes mérőprogram nem követeli meg tökéletes szinkronban lévő, csúszásmentes órák alkalmazását, ezért el kellett vetnünk annak lehetőségét, hogy mindenki-mindenkinek módon Dimes alkalmazások között megvalósítsuk a tomográfias mérést. A mérésben szükséges megbízható órareferenciák használata, amit az Etomic számítógépek biztosítanak. Azaz minden mérni képes Dimes alkalmazás felől egy, az Etomic végpontok irányába húzódó fa szegmenseit tudjuk csak felöltöztetni becsült sorhossz statisztikákkal. Az Etomic óráit referenciának használva korrigálni tudjuk a Dimes mérőprogramot futtató gép órájának csúszását. Abszolút pontos órabeállításra nincs lehetőségünk. A csúszás korrekciója után a Dimes időbélyegeit egy konstans eltolás erejéig tekinthetjük pontosnak, ami a tomográfias kiértékelésnél elegendő. A tervezésnél figyelembe vettük, hogy a Dimes program alacsony prioritással fut a gazdaszámítógépen és csekély forgalmat kelthet, ezért nem küldhető vele hosszú csomagvonal. Ennek fényében egy mintázatot négy csomag alkot úgy, hogy az első csomagpár célja ugyanaz az Etomic végpont, míg a második pár címzettje egy másik Etomic végpont. Erre azért van szükség, hogy a páronként érkező cso-



3.12. ábra. A Dimes mérőalkalmazás kiegészítő **packettrain** modul, aminek segítségével a felhasználó által felparaméterezett csomagvonal indítható periodikusan a hálózatba. A mintázatot azonos (P) méretű csomagok építik fel, melyek címzetjei (IP) és az előző csomaghoz viszonyított kiküldési idejei (δ) egyenként beállíthatók. A mérés során a modul N darab mintát küld dt nagyságú szüneteket beékelve a hálózatba. A próbacsomagok egyedi sorszámmal bírnak és az adatmező tartalmazza a küldő végpont időbélyegét valamint a Dimes alkalmazás azonosítóját.

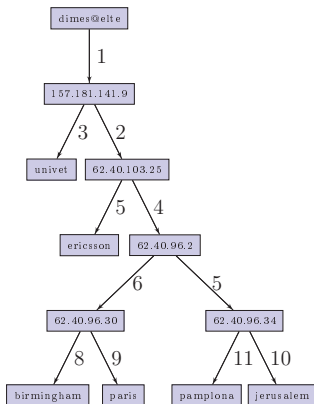


3.13. ábra. A `dimes@elte` \rightarrow `jerusalem.etomic.org` irányú hálózati késleltetés és adatokat az óra csúszását figyelembe véve korrigáltuk és konstans eltolás erejéig meghatároztuk. Az idősor határozott minimummal rendelkezik, ami megfelel annak az esetnek, hogy a próbacsomag a hálózaton sorbanállási késleltetés nélkül haladt keresztül. A felfelé szóró pontok az egymás utáni sorokban eltolított várakozásból adódnak. Az adatok felső 1%-át kiszűrve a valószínűségi sűrűséget is ábrázoljuk.

magok korrelációs tulajdonságait vizsgálva megbízhatósági mutatónk legyen arra, mennyire tekinthető az eredeti minta egyszerre kiküldöttnek, és a közbülső két csomag adatait használjuk az Y topológiára kidolgozott tomográfias becslésben.

Az elvégzett referencia méréseinkből a 2005. július 12-én begyűjtött adatokat mutatjuk be a dolgozatban. A mérésben a Komplex Rendszerek Fizikája Tanszéken telepített fejlesztői Dimes alkalmazást használtuk küldőként, és a következő hat Etomic végpont felé küldtük párosával próbáinkat: $\{univet, ericsson, birmingham, pamplona, paris, jerusalem\}.etomic.org$. Minden végpont pár felé $N = 10000$ darab mintavonatot küldtünk a hálózatba. A $P = 40$ byte méretű, szorosan illeszkedő ($\delta = 0$) csomagokból felépített csomagvonatok közé $dt = 10000$ ms várakozási időket iktattunk, elkerülendő a hálózat túlterhelését.

A mérés során begyűjtött idősorok drifftel terheltek, ami a küldő órájának csúszása miatt mérési műtermék. Ezt a hibát könnyen korrigálhatjuk. Ha feltételezzük, hogy a mérés során számos próba üres sorokon haladt keresztül a hálózatban, akkor minden idősornak határozott minimummal kell rendelkeznie, ami a 3.1. képlet alapján a vonalakra jellemző *propagációs* idők valamint a belső eszközök kiszolgálási idejeiből adódik össze, azaz konstans. Ezen elvárás mellett az adatok *felfelé* szóródnak. Amennyiben csomagok indulását és érkezését megfigyelő órák bármelyike csúszik, az idősor konstans minimuma, lineáris alsó burkolónak torzul a valódi adatokon. Esetünkben is pont ez a helyzet. A referenciaórák mindegyikéhez az idősorok alsó burkolójára egyenest illesztettünk, majd a meredekségek átlagát tekintettük a küldő óra ráta hibájának. A konkrét mérésben az óra csúszása 19 ± 1 $\mu s/s$ -nek adódott, amit figyelembe véve korrigáltuk az idősorokat, ahogy az a 3.13. ábrán is látható.



(a) Feltárt hálózat

Él	$E[x]$	Hiba	$\sigma[x]$	Hiba
1	23,446	2,738	35,044	4,652
2	6,976	3,883	5,052	2,048
3	32,196	2,779	70,856	2,678
4	6,693	4,016	17,467	10,557
5	83,373	5,730	48,426	1,373
6	1,802	0,707	6,059	4,215
7	8,273	1,452	5,441	0,981
8	43,575	—	30,752	—
9	30,866	—	37,188	—
10	309,853	—	294,230	—
11	92,723	—	136,885	—

(b) Sorhossz statisztikák

3.14. ábra. A dimes@elte mérőprogram 6 Etoic végponttal közösen végezte a tomográfiás mérést. (a) A **traceroute** mérések alapján előállított feszítőhálózat. (b) A becsléssel feltárt sorhosszeloszlások várható értékei, szórásai és hibái $\mu\sigma$ egységekben.

Az 5.2. alfejezet módszereit követve elvégeztük a sorhossztomográfiát, melynek segítségével a feltárt 49 útvonalválasztó közül 5 elágazási pontot azonosítottunk a mérésben, és 56 átfedő szegmensre adhattunk sorhossz eloszlás becslést. Az átfedő szegmenseket átlagolva 11 egyedi szegmensre jellemző sorhossz statisztikát tártunk fel, az eredmények részletesen a 3.14. ábrán láthatóak.

3.6. Diszkusszió

Célul tűztük ki, hogy a valódi hálózatban megteremtjük annak a lehetőségét, hogy méréseken alapuló, megbízható torlódási térképet állítsunk elő. Ehhez elengedhetetlen volt a megfelelő mérőhálózat megtervezése és megalkotása. A fejezetben bemutatunk a nemzetközi összefogásban megépített Etoic mérőrendszert és bemutattuk azokat a mérési módszereket, amiket torlódással kapcsolatos méréseinkben használunk. Betekintést nyújtottunk abba, hogyan tudjuk más meglévő mérőrendszerek segítségével szélesebb körbe kiterjeszteni kísérleteinket, azok milyen technikai kihívásokat támasztottak. A fent bemutatott mérések gyakorlatilag topológia feltáró és csatolt hálózati késleltetés méréseket valósítanak meg, melyeket az 5. fejezet tomográfiás becslési módszereivel dolgozunk fel, hogy megrajzolhassuk a hálózat torlódási térképét.

4. fejezet

Hálózatmérési virtuális obszervatórium

4.1. Motiváció

A tudományos kísérletekből származó naprakész és archív mérési adatok közhatalossága, a tudomány számos területére jó hatással volt. Az információáramlás felgyorsulása serkenti a kutatók munkáját, modelljeik összevetése könnyebben megvalósítható, gyakrabban születnek új észrevételek. A tudomány más területén meghonosított elveket követve kutatócsoportunk felkarolta azt az irányzatot, hogy a hálózati mérések területén is terjessze és felépítse a hálózati mérések virtuális obszervatóriumának egy lehetséges prototípusát. Az obszervatórium életre hívásával legfőbb célunk az volt, hogy a már elérhető mérési adatok és a különböző mértékek becslési eredményeit megbízható referenciaként közzé téve, azok használatára buzdítsuk a hálózati forgalom kutatásával foglalkozó szakembereket. A tervezés és építés során igyekeztünk megteremteni annak lehetőségét, hogy más kutatócsoport is beilleszthesse az Internet hálózattal kapcsolatos adatait és eredményeit a kidolgozott keretrendszerbe. Munkánk során folyamatosan foglalkozunk az adatok hatékony tárolásának illetve szabványos megjelenítésének kérdéskörével is.

A mondást, mely szerint „*a tudás hatalom*”, napjainkban kicsit át kell értelmeznünk. Mára talán az igazi hatalmat az információ megfelelő közzé tétele jelenti, hiszen a nyilvános adatok és a hozzáférhető algoritmusok alapján működő különböző kiértékelési módszerek hozzásegítik a tudományt a gyorsabb és hatékonyabb fejlődéshez. Ezáltal a tudományos felfedezések üteme nő. Említésre méltó, hogy a Humán Genom Projekt [45] keretén belül, az elsők között építettek fel tudományos adatbázisokat. A tudomány számára a munka nem állhat meg az adatbázisok felépítésével és feltöltésével, hiszen a közzétett adatok más kutatócsoportok munkáját is elősegítik. A fenti példát alapul véve, az emberi génnek térképét elkészítő kutatócsoportoktól független csoportok vizsgálják a külön-

böző fajok biokémiai reakciói mögött megbúvó komplex reakcióhálózatokat. Azok a kutatók, akik egy kiválasztott fehérje evolúcióját vagy annak fiziológia tulajdonságait akarják megvizsgálni, hasznos a két fenti kutatás nyersadatait együtt elérni, független adatbázisaikat együtt kezelni, hiszen sokszor elkerülhetetlen a legelemibb mérési adatokig visszanyúlni, mert nem elegendő valamilyen kiértékelte eredményekre támaszkodni.

Elsőként a csillagászat tudományában [46, 47] sikerült legyőzni azt a technikai akadályt, hogy számos különböző észlelési munka adatait kiszolgáló adatbázisokat kötegeljenek. Ezen munkát végző szakemberek egyik fontos vívmánya, hogy a több adatbázison átnyúló objektum keresés megvalósult. Az ún. *cone search* eljárás használatával az égbolt egy meghatározott területére vonatkozó összes, az adatbázisokban megtalálható mérési adat egybegyűjtve kerül a kutató rendelkezésére.

Az alábbi pontokban egybegyűjtjük a hálózati mérések virtuális obszervatóriumának legfontosabb céljait:

- Módszer bevezetése a referencia mérési adatok nyilvános elérhetőségére;
- A mérésekből gyűjtött nyersadatok különböző kiértékeléseiből származó eredmények megosztása;
- Külső, szabványosított nyersadatok kiértékelésének biztosítása;
- Az adatok manipulálásához könnyen kezelhető felhasználói felület beüzemelése;
- *Web-szervízek*en keresztül csatlakozási felület biztosítása a hálózati alkalmazásoknak;
- Szabványos csatlakozási pont biztosítása más kutatócsoportok archívumai felé;
- A hálózatot jellemző különböző mértékek ábrázolásának kérdésével kapcsolatos megbeszélések inspirálása;
- Sémajavaslatot nyújtani az újonnan felépülő hálózatmérési adatbázisok tervezésénél.

4.2. Elérhető referencia nyersadatmegosztók

A hálózati mérésekkel foglalkozó eddigi munkák közül fontos bemutatnunk pár példát. Az Etomic rendszer nyersadatait kiszolgáló virtuális obszervatórium felépítésénél figyelembe vettük ezen törekvések ötleteit és technológiai eredményeit.

Az amerikai székhelyű Caida 2002-ben hozzáfogott ahhoz, hogy aktív és passzív mérésekből származó adatait megfelelően feldolgozza és pár releváns tulajdonság alapján, jól kereshető rendszerbe szervezze. Felépítette az Internetmérések Nyersadat-katalógusát [28] (*DatCat*). A kidolgozott rendszerben mára a Caidától független csoportok mérési nyersadatai is katalogizálhatóvá váltak. A megvalósított keretrendszerben csak címkékből álló, az ún. *meta-adatok* tárolása valósul meg, a valódi nyersadatok kötegeire hálózati hivatkozást (*URL*-t) biztosít. A közvetett módon elérhető nyersadatok formátumára nincs semmilyen megszorítás, a kutató feladata az esetlegesen különböző forrás eltérő formátumait harmonizálni, ha az adatok együttesével kíván dolgozni. A DatCat-ban fellelhető nyersadatok között passzív csomagszintű naplók, *traceroute* mérési fájlok, útvonalválasztási táblázatok¹, hálózati vírussterjedési térképek, stb. szerepelnek.

A DatCat analógiájára futott a Mome Projekt [48] Európai Unió IST Program keretében. E munka fontos újításai közé tartozik, hogy a szabványos reprezentáció felé megtették az első lépéseket. Így a *traceroute* mérések formátumára kritériumrendszert alkottak [49], valamint a mérési nyersadatok kiértékelési algoritmusaira megosztására és futtatására lehetőséget nyújtanak.

A MAWI munkacsoport [50], ami 1999. óta a japán *WIDE* gerincvonal számos pontján rendszeres passzív méréseket végez, a megfelelő módon anonimizált mérési naplókat valamint az ezalapján automatikusan kiértékelte, a forgalmat jellemző statisztikát teszi közzé.

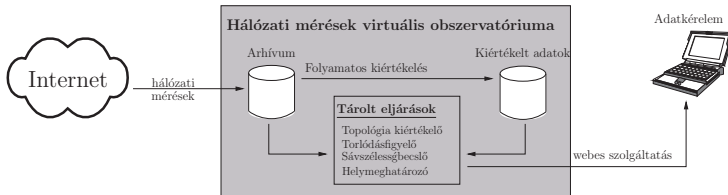
Az ímént bemutatott mérési nyersadatkiszolgáló rendszerek megragadnak saját adataik publikálásánál illetve több nyersadat forrás esetén, megelégednek az adatok szabványos feldolgozásával. A következőkben bemutatott virtuális obszervatórium megpróbál ezen túllépni. Keretrendszert javasunk, mely segítségével különböző források megoszthatják adataikat és létrehoztunk egy adatbázisbeli sémarendszert, amelynek segítségével a csomag alapú mérések nyersadatai elegáns és hatékony módon tárolhatók.

4.3. Az obszervatórium prototípusa

A komplex rendszerek működésének megértéséhez komplex modelleket alkotunk. A modellek előrejelző képességét statisztikailag reprezentánsnak tekinthető minta segítségével tesztelhetjük megbízhatóan. Például a hálózati forgalom és a topológia dinamikájának térbeli és időbeli viselkedését leíró új modellekhez nagy lefedettségű, hosszú múltra visszatekintő megfigyelési adatokra van szükségünk. Az átfogóbb vizsgálatok reményében saját méréseinkből származó adatainkat ki kell egészítsük más hozzáférhető hálózati mérési forrással. A különböző, egymástól független kutató csoportok által szolgáltatott nyersadatok felvetik az adatok összevethető ábrázolásának és hatékony tárolásának problémáját. Az ábrázolás

¹ *BGP* – Border Gateway Protocol táblák

tekintetében megfelelő, mindenki által elfogadott szabványos adatmodellek kifejlesztésére csak nemzetközi összefogásban remélhetünk megoldást. A hatékony tárolás megvalósítására a megoldást, esetünkben is, jól szervezett relációs adatbázisok üzemeltetése jelenti, melyben a sémarendszer hűen illeszkedik a hálózati mérések jellegéhez. Természetesen a sémarendszer részletes tárgyalása, ami szép technikai kihívás, nem képi a dolgozat részét.



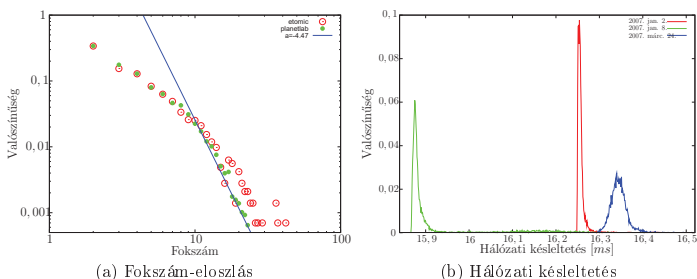
4.1. ábra. A hálózati mérések obszervatóriumának felépítési vázlata. A hálózati mérések paraméterei és nyersadatai az Archívumban tárolódnak. Az adatokat rendszeresen kiértékelő, tárolt eljárás formájában megvalósított, statisztikai becslők szintén az adatbázisba töltik az eredményeket. Adatkérelem esetén a nyers és a kiértékelt adatok webes szolgáltatásokon keresztül érhetők el.

A 4.1. ábrán tömören vázoljuk a megvalósított obszervatórium belső felépítését. Az obszervatórium lényeges eleme az adatok tárolására üzembe helyezett adatbázis. Az Etomic mérőhálózaton rendszeresen végzünk hálózati méréseket, mint például sorhossz tomográfiás, topológiafeltáró, geográfiai lokációs és sávszélességbecslő mérések. Mérési adatainkat az adatbázisban archiváljuk. Azt az elvet követjük, hogy a mérés során begyűjtött nyersadat minden lehetséges részletét tároljuk, noha így elkerülhetetlenül gyorsan hízó adattáblákat is létrehozunk. A „mindent markolni kívánó” elv mögött meghúzódó érvelés a következő. Egyfelől elképzelhető, hogy az eddig bevezetett becsléseken felül születnek új, finomabb, vagy teljesen más mértékeket vizsgáló kiértékelő módszerek, melyek csak akkor képesek az archív referencia adatokat feldolgozni, ha a mérőforgalom minden lehetséges részlete reprodukálható. Másfelől evidens, hogy a ma elvégzett Internet mérések mikroszkópikus tulajdonságai a jövőben szükségszerűen teljesen más képet festenek. Csak akkor remélhetjük helyesen megérteni a tapasztalt apró eltéréseket, ha a mérés a lehető legpontosabban van az adatbázisba beleképezve. Harmadrészt nem ismerhetjük, hogy a nyersadatok milyen kombinációjában állnak elő a nagyobb időskálán releváns statisztikák. A jövőben felmerülő ötletek szabad kipróbálása is megköveteli a részletekbe menő adattárolást. A részletekbe menő tárolásra példaként említjük tomográfiás méréseink nyersadatait. Itt sem elegendő pusztán az idősorok és a topológia tárolása, szükséges a próbaforgalom csomagszintű részletes leírása azért, hogy ezeket a mérési adatokat egy ma még

nem létező kiértékelő eljárás felhasználhassa.

A részletes tárolásból adódik, hogy a nagy méretű adattáblák kezelése fontos probléma, amivel foglalkoznunk kell. Az adatbázis felé eddig megfogalmazott jellemző kérdésekre alapozva kutatjuk a táblázatok keresést gyorsító hatékony felindexelését.

Az obszervatórium részét képezi számos olyan függvény, melynek segítségével sokszor előforduló bonyolultabb lekérdezéseket lehet átláthatóan megfogalmazni. A tárolt függvények és eljárások között megvalósítottunk hasznos statisztikai eszközöket, melyek segítségével például hisztogramot szerkeszthetünk idősorokhoz, vagy az eloszlások különböző momantumait értékelhetjük ki. Ezek a függvények és eljárások web-szolgáltatásokon keresztül is rendelkezésre állnak, így lehetőség nyílik arra, hogy könnyebben illeszthessük a rendszerhez az adatokat elemző külső programokat.



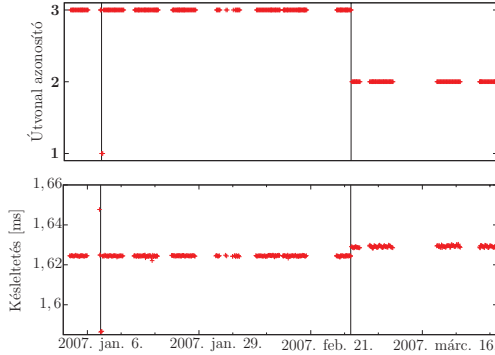
4.2. ábra. Az obszervatórium segítségével megválaszolható kérdésekre két eklatáns példa. (a) Az adatbázisba eddig begyűjtött *traceroute* mérések alapján előállított hálózat fokszám eloszlásának vizsgálata. A mérésben részt vevő végpontok alapján két halmazt vizsgáltunk meg. Az egyikben csak az Etomic rendszer végpontjai között, a másikban a Planetlab infrastruktúra ágensei között feszülő irányítatlannak tekintett hálózat fokszámát ábrázoljuk. (b) Az archiv idősorok hosszú távú viselkedésének tanulmányozása is lehetséges. A *univet.etomic.org* és az *magdeburg.etomic.org* végpontok között mért hálózati késleltetésének valószínűsítését látjuk három különböző időpontban.

Munkánk gyakorlati használhatóságának alátámasztására, az alábbi kérdésekre keressük a választ a megvalósított obszervatórium segítségével.

Az eddigi méréseink során érintett hálózat topológiája milyen fokszám eloszlást követ? Az eddig archivált *traceroute* alapú útvonalfeltáró mérések nyersadataiból előállított hálózati topológia, 4.2a. ábrán látható fokszám eloszlása szerint, jó közelítéssel hatványfüggvény szerint esik, melynek kitevője $\alpha = -4,47$ körüli. A

kitevő nem mutat lényeges eltérést aszerint, hogy kisebb, csak az Etomic végpontok között feszülő hálózatot vizsgálunk, vagy egy nagyságrendekkel több elemből felépülő hálózatot, ami a Planetlab nódusok között található.

Tapasztalunk-e valami érdekeset a különböző mérésen átnyúló hálózati késleltetés idősorában? A 4.2b. ábra alátámasztja, hogy az obszervatórium segítségével az archív idősorok hosszú távú viselkedése tanulmányozható. Két végpont között mért hálózati késleltetésének sűrűségfüggvényét mutatjuk meg három különböző 2007. év elején végzett mérésre. A kiragadott példa azt sugallja, hogy a különböző mérés során a mérőcsomagok különböző útvonalon haladtak keresztül. A január elején tapasztalható késleltetésnél kisebb hálózati késleltetés volt tapasztalható pár napig a hó végén, majd a márciusi adatokon újabb eltolódást tapasztalunk. Nagy valószínűséggel megváltozott az mérőcsomagok által bejárt útvonal.



4.3. ábra. Az *univet.etomic.org* és a *magdeburg.etomic.org* végpontok között megfigyelhető az útvonal és a késleltetés között rejlő természetes kapcsolat. A kiragadott időszakban 3 különböző útvonal típust mértünk melyekkel jól korrelál a hálózaton tapasztalható *propagációs* és kiszolgálási idő.

Találhatunk megbízható evidenciát útvonalváltásra az idősorok elemzése alapján? Ez a kérdés az adatok átfogóbb együttes tanulmányozását követeli meg, melyre kiválóan alkalmas az obszervatóriumunk. Az eddigi mérési adatokat megvizsgálva a mérőpontpárokat két csoportra bonthatjuk. Vannak olyan útvonalak, ahol nehéz felállítani egyértelmű kapcsolatot a hálózati késleltetés és a feltárt útvonal között. Ezekben az esetekben nagy valószínűséggel úgy módosul az útvonal belső felépítése, hogy a szakaszokat alkotó fizikai eszközök cseréje nem vonzza a címzések cseréjét. A másik csoportba azok a párok tartoznak, ahol könnyen kimutatható a kapcsolat. A 4.2b. ábra adatait szolgáltató budapesti és magdeburgi

Útvonal	Elemsszám	$\min \tau [ms]$	$E[\tau] [ms]$	$\sigma[\tau] [\mu s]$	Úthossz
1	3	15,86	–	–	14
2	125	16,28	16,29	4,4	15
3	267	16,22	16,25	14,5	13

4.1. táblázat. Az *univet.etomic.org* és a *magdeburg.etomic.org* végpontok között megfigyelt 3 különböző útvonal késleltetés statisztikái. Jól látható, hogy egy fizikailag állandó útvonalra leszűkítve kicsi, $\approx 10 \mu s$ nagyságrendű ingadozásokat tapasztalunk.

végpont között jól kimutatható ez a kapcsolat. A 4.3. ábrán megvizsgáljuk időben hogyan viselkedik az említett két végpont között a minták késleltetése. Minden mérési nyersadatköteghez meghatároztuk a hálózati késleltetés minimumát, azaz a *propagációs* és a kiszolgálási időkből származó állandó késleltetést. Az adatokban 2007. január 8-án és 2007. március 1-én egy-egy éles ugrást találunk. Erre a mérési időszakra szorítkozva az útvonalfeltárásából származó adatokat összevetve a késleltetésekkel, időben ugyanakkor tapasztalunk változást. Az **1**, **2** és **3** jelű utak mikroszkopikus szerkezete eltér, a hálózat magjában átrendeződés figyelhető meg. A különböző utakhoz tartozó idősorokat megvizsgálva hasznos tanulságot vonhatunk le. Elvárásainknak megfelelően minden stabil és üres útvonalhoz állandó késleltetés tartozik. A vizsgálat alapján empirikus küszöbszámot rendelkezhetünk ahhoz, hogy mit tekintünk állandónak. A 4.1. táblázat alapján a késleltetések szórása a $10 \mu s$ nagyságrendbe esik. Ezt a tudást felhasználhatjuk arra, hogy két tetszőleges mérésből származó idősort összevetve becslést adjunk arra, azonos útvonalon valósult-e meg a mérés. E kérdéskört tovább vizsgálva útvonal-tomográfias eljárásokat lehet kidolgozni.

4.4. Diszkusszió

Ebben a fejezetben felvázoltuk annak jelentőségét, hogy a tudomány más területén jól bevált utat követve, a hálózati mérések tekintetében is szükség van a mérési tapasztalatok közkinccsé tételére. Ez a törekvés a mindenkor mérési adatok és mérési körülmények tárolásán felül a folyamatosan bővülő statisztikai kiértékelő eljárások megosztását is jelenti. Érintőlegesen felvázoltuk az Eto-mic mérőrendszer mellé felállított adatbázis működését, melyben eddigi méréseinket és a belőlük következő eredményeket szervezzük és tároljuk. A mérési adatok historikus korrekt tárolásának segítségével a nagyobb időskálán bekövetkező változások megfigyelésére és elemzésére nyílik mód. A keretrendszer lehetőséget nyújt, hogy web-szolgáltatásokon keresztül is kiszolgálja az adatokat, ezzel megteremtve intelligens alkalmazások tervezésének lehetőségét. A virtuális obszervatórium problémáin, az Európai Unió VII. Keretprogramjában meghirdetett Mo-

ment pályázat [51] keretében dolgozunk tovább. A pályázat legfőbb célkitűzése, hogy különböző kutatási csoportok meglévő adatbázisait illetve on-line elérhető mérési alkalmazásait egységes felület mögé szervezze úgy, hogy a hálózatkutató kéréseire szabványos, több forrásból származó aggregált adatot nyerhessen ki. A mérési adatok elérhető történelmi archívumai elengedhetetlenül segítik az Internet időfejlődését kutató adatbányászok munkáját.

5. fejezet

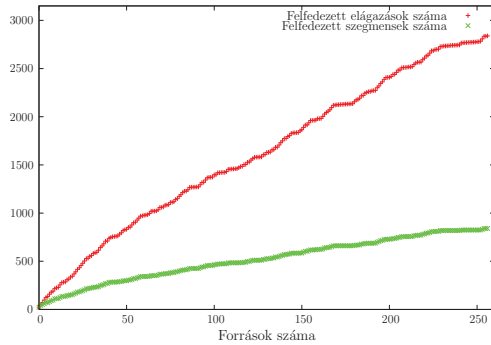
A hálózati pufferek sorhosszainak becslése tomográfiás módszerekkel

5.1. Motiváció

A Bevezetőben tárgyaltak szerint az Internet felépítését és forgalmát tekintve, komplex rendszer. A hálózat működését és állapotát leíró gyors és megbízható becslések kifejlesztése elengedhetetlen feladat, ha hatékonyabb forgalomvezérlő protokollokkal tervezzük felváltani a most használatban levőket, ha meg akarjuk oldani a torlódásokat kikerülő automatikus útvonal átrendezés problémáját, illetve ha egyszerűen fel akarjuk tární a hálózat gyenge vagy támadás alatt lévő helyeit. Az Interneten ma legelterjedtebben használt csomagküldő protokollvariánsok szabályzási döntéseihez legfontosabb állapotjelzők a hálózatban tapasztalható csomagvesztési valószínűségek és a késleltetések. Az állapotjelzők, mint a vesztési valószínűség és a késleltetés helyi monitorozó mérése, az adatok begyűjtése és elemzése nagy skálán megoldhatatlan feladat, többek között az érintett különböző alhálózatok eltérő gazdasági érdekeltségei miatt. Bár a hálózat belsejében üzemelő eszközök gyűjthetnek forgalmi statisztikákat, az adatok tárolása és elemzése az eszközök drága módosításait vonná maga után eszközfajlesztés és karbantartás oldalról egyaránt. A forgalomról nyert információ továbbítása, esetlegesen központosított tárolása roppant nagy extra hálózati forgalmat eredményezne.

Az Internet tudományos kutatásában egyre nagyobb tért hódítanak a hálózattomográfiás eljárások, melyek kiválóan alkalmazhatóak a szabályozatlan fejlődésű heterogén felépítésű hálózatok tulajdonságainak becsléseken alapuló tanulmányozásában. Méréstechnikai oldalról megközelítve a problémát, minden hálózattomográfia kulcs gondolata, hogy aktív méréseket végzünk, – lásd a 3.3.2. alfejezet bemutatott technikáit, – melyek a hálózat belsejében működő eszközök speciális közreműködését nem igénylik, és amik így összességében kisebb plusz forgalmat generálnak. A mérések közvetve olyan fizikai mértékek meghatározására

irányulnak, amiket közvetlenül nem tudunk leolvasni, például a hálózat belsejében tapasztalható egyedi késleltetések feltárása. A rejtett információ felderítésére a statisztika dedukciós módszereit alkalmazzuk. Vardi [52] éppen azért vezette be a *hálózattomográfia* fogalmát, hogy kihangsúlyozza az orvoslásban is használatos tomográfia eljárásaival a rokonságot. Mivel a hálózattomográfias becslések a végpontokon gyűjtött statisztikákon nyugszanak, ezért a mérési infrastruktúra jól skálázó, eszközigényét és karbantartási költségeit tekintve olcsó eljárásnak számít.



5.1. ábra. A mérőpontok számának függvényében felfedezhető elágazási pontok és szegmensek számának változása a gyakorlatban. A topológiát feltáró mérésben a világon elszórtan üzemelő 257 Dimes mérőprogramot használtunk és 15 Etomic végpont felé mértünk.

A hálózattomográfia legfontosabb problémái közé tartozik, hogy számos útvonalon megvalósított aktív végpontmérések statisztikáira alapozva dekomponáljuk a hálózat belsejében levő szegmensekre jellemző késleltetések [53, 54] illetve veszteségi valószínűségek mértékeit [55, 56]. A tomográfias mérésekben általában egy végpontból a végpontok egy halmazába mérünk, és ebben az esetben a forrás és a nyelők között feszülő logikai fa élére határozzuk meg a késleltetés vagy a vesztesés statisztikáját. A feltárható logikai fa elágazási pontjai megfelelnek azoknak az útvonalválasztó egységeknek, ahol a próbaforgalom csomagjai különválnak. Növelve a tomográfias mérésben részt vevő források és nyelők számát, a hálózat egyre nagyobb részére képes az eljárás állapotbecslést szolgáltatni. A forrás és nyelő végpontok megegyező halmazát feltételezve 3.3.2. szakasz a 3.3. ábrája alapján extrapolálhatunk a feltárható hálózat méretére. Aszimmetrikus forrás és cél végponthalmaz esetén, – ilyen például a 3.5. fejezetben említett Dimes-os konfiguráció, – az 5.1. ábra ad becslést a feltérképezhető hálózat nagyságáról.

A hálózattomográfiás technikákat eredendően olyan hálózatokra fejlesztették ki, ahol az útvonalválasztó eszközök képesek *multicast* forgalomra, azaz képesek olyan különleges csomagok kezelésére, amikben egyszerre több cím is fel van tüntetve. Amennyiben az útvonalválasztóba érkező *multicast* csomagban található címek egy része az egyik kimenő vonala felé érhető el, míg a többi cím más vonalon, akkor az eszköz annyi új csomagot készít, ahány irányban szükséges tovább küldeni az eredeti csomag tartalmát. A *multicast* tomográfiás [57, 58, 59, 60] mérések a hálózati elemek együttműködését követelik meg, ami általában az Interneten nem teljesül. Ezért a becslési módszereket továbbfejlesztették a hagyományos címzésű csomagokra [55, 53, 56]. Érdekességként megemlítjük, hogy több forrás módszerrel találunk példát a [61, 62, 63] munkákban.

Az általunk használt *unicast* hálózattomográfia létjogosultsága mögött húzódó fizikai gondolat a következő. A forrásként működő mérési végpont nagyon szorosan egymás mögött elhelyezkedő, ún. *back-to-back* csomagokból összeállított vonatot küld a hálózatba úgy, hogy az egyes csomagok címzettjei különböznek. A mintavonat hossza szerint, – megvalósításunk időrendjét követve, – megkülönböztethetünk két mérési alternatívát. Legegyszerűbb esetben egy forrás és két végpont szerepel a mérésben, ezt Y-tomográfiának nevezzük. Amennyiben több, mint két végpont felé történik a tomográfiás mérés, fákra kidolgozott tomográfiás becslésről beszélünk. A szorosan illeszkedő mérőcsomagok mindaddig ugyanazt az utat járják be a forrást elhagyva, amíg címeik szerint a hálózat belsejében szét nem válnak. Ezalatt ugyanazokon a hálózati eszközökön haladnak keresztül koherensen, ugyanazokat a sorokat állják végig. Gyakorlatilag így a hálózat ugyanazon állapotát tapogatják le. E munkahipotézis lehet az alapja annak, hogy a minták hálózati késleltetés és idősorai korrelálttá válnak. A mintákban megtapasztható korrelációnak köszönhetően ez a mérés gyakorlatilag nagyon hasonló az Interneten csak elvben kivitelezhető ideálisabb *multicast* mérésre, és ennek a tulajdonságnak köszönhető az is, hogy a végpont-végpont adatokból a hálózat belső tulajdonságaira következtethetünk.

A 3. fejezet 3.1. képlete alapján látszik, hogy a hálózati késleltetés nemnegatív komponensek összegeként áll elő minden útvonalra, mely tagokból a $\Sigma(\tau + t_x)$ állandónak tekinthető, amihez egy időben ingadozó, a sorban állásokból származó Σx idő adódik hozzá. A következő fejezetekben a sorbanállási időket leíró valószínűség-sűrűségek tomográfiás becslési eljárásait mutatjuk be. Majd az Etomic rendszerben végzett mérésekből mutatunk példákat, és ismertetjük az eredményeket, mely a hálózat belső torlódási állapotáról hordoz érdemi térbeli információt. Az Interneten végzett kísérleteket a nap különböző szakában megismételve a hálózat időbeli fejlődését is nyomon követhetjük, amiről egy animációs filmet mellékelünk. Bízunk abban, hogy az Interneten rendszeresen végzett nagy skálájú tomográfiás méréseink és az adatok feldolgozása és közzé tétele hozzásegít a hálózat késleltetéseit vizsgáló egyre realiztikusabb modellek megalkotásához, ami érdekes aktív kutatási terület.

Annak érdekében, hogy a hálózattomográfiás eljárások a passzív mérési tech-

nikák eredményeivel felvegyék a versenyt, törekedni kell a nagy számításigényű algoritmusok finomítására, melyre mi is bevezettünk pár ötletet.

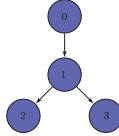
5.2. Tomografikus sorhosszeloszlás-becslés Y-alakú hálózatban

Ebben az alfejezetben azt a statisztikai becslést mutatjuk be, aminek segítségével a csomagok az útvonalválasztó egységek puffereiben eltöltött várakozási idejeit tárjuk fel. A módszer arra az elemi hálózatra van kidolgozva, amiben egy forrás és két nyelő szerepel. A próbacsomagok együtt haladnak a forrástól az elágazási pontig, majd két felé válva Y-alakú utakat járnak be. Annak érdekében, hogy becslést adhassunk összetettebb hálózatra is, azt az összes lehetséges Y-formával fedjük le. Kísérletileg a csomagpárokat felváltva különbözően előállított célpontpár felé küldjük. A hálózat Y-okkal történő lefedése következtében az utaknak lesznek teljesen átlapoló és csak részben átfedő szakaszai. Feltételezzük, hogy a mérés teljes idejére állandósult állapotban van a hálózat, az egyedi szegmensek sorait leíró valószínűségi eloszlások nem változnak a mérés ideje alatt. Ezt a gyakorlatban úgy tudjuk biztosítani, hogy a különböző célpár kombinációk mintáit váltakozva utaztatjuk, így az átfedő útszakaszokról ugyanazt az állapotinformációt gyűjtjük. Az egyedi Y-kiértékelés eredményeit a teljesen átfedő szakaszokon konzisztencia vizsgálatra használhatjuk, hiszen ugyanazt az állapotot várjuk el független mérési mintáink alapján. A részben átfedő szakaszok segítségével, inverz konvolúciós eljárással a két eredeti szegmens három rész-szegmensre bontható, ezzel a teljes hálózat finomabb felbontására nyílik lehetőségünk.

A módszer szépsége, hogy a becslés képletei zárt alakba hozhatók, melyek kiértékeléséhez számítógépes numerikus módszereket alkalmazunk. A mérési módszer gyengéje, hogy növelve a végpontok halmazát, exponenciálisan egyre több mintára van szükség, ami rontja a skálázás esélyeit. A több végpontot tartalmazó mérésekben megnövekedett mintaszám elnyújtja a mérés időtartamát, ami határesetben rontja a rendszer állandósult állapotára kirótt feltételezésünket. A mérés ideje alatt jelentősen megváltozhat a hálózat dinamikája. A minták sűrűbb küldése sem segít, mert eldugítva a hálózatot a rendszert kibillentjük a mérni kívánt állapotából.

5.2.1. A becslési módszer

Először az elemi tomográfias sorhosszbecslés módszerét mutatjuk be, ami az egy forrás, két nyelő végpont és egy elágazási pont között feszülő Y-alakú útvonalnak felel meg, lásd az 5.2. ábrán. Méréstechnikailag fontos megjegyeznünk, hogy az ábra egyszerűsítés eredményeként előálló feszítőfa. Az itt feltüntetett utak a valós méréseinkben nem egyetlen vonalat jelentenek, hanem vonalak és útvonalválasztók olyan folytonos láncolatát, melyet már nem tudunk kisebb elemekre



5.2. ábra. A legegyszerűbb hálózati elem az Y-alakú elágazás. A 0-val jelzett forrás végpont szorosan illeszkedő csomaggpárokat küld a 2 és a 3 célpontba. A csomagok hálózati késleltetéseit $(Y_2, Y_3)^{(i)}$ -vel jelöljük az i -ik mintához tartozó mérési adatokat.

bontani.

Az Y-alakot megvalósító elemi tomográfias mérésben a forrás (0) szorosan illeszkedő csomaggpárokat küld a célpontok (2 és 3) felé, amik az Y szarán együtt haladnak, majd az elágazási pontban (1) szétválnak. A csomagok hálózati késleltetését mérjük, amiből statisztikai módszerek segítségével következtetünk vissza a hálózat egyes szegmenseinek egyedi késleltetésére.

Végezzünk N mérést, melyből az i -ik csomaggpárra (Y_2^i, Y_3^i) késleltetéseket kapunk, ahol az alsó indexek a célpontokat jelölik. Ebben a modellben tekintsük a belső késleltetéseket és a hálózati késleltetéseket valószínűségi változóknak, amiket vektorokba rendezve rendre $\underline{\mathbf{x}} = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ és $\underline{\mathbf{y}} = (\mathbf{y}_2, \mathbf{y}_3)$ jelölünk. A késleltetések alsó indexeit úgy választjuk meg, hogy a hozzátartozó irányított élek érkezési oldalán található csúcs jelével egyezzenek meg.

A rejtett és a megfigyelhető valószínűségi változók között természetes, a topológiától függő lineáris összefüggés teremt kapcsolatot:

$$\mathbf{y}_2 = \mathbf{x}_1 + \mathbf{x}_2, \quad \mathbf{y}_3 = \mathbf{x}_1 + \frac{P}{C} + \mathbf{x}_3, \quad (5.1)$$

ahol második képletben megjelenő gyakorlatilag állandónak tekinthető P/C tag azt írja le, hogy a 2-es célba menő P méretű csomag kiszolgálási ideje feltartja a 3-as célba haladó próbát.

Az 5.1. képleteket alkalmazva a mérési eredményekre, írjuk hogy

$$Y_2^i = X_1^i + X_2^i, \quad Y_3^i = X_1^i + \frac{P}{C} + X_3^i, \quad \forall i \in \{1, 2, \dots, N\},$$

amiből jól látszik, hogy minden mérési eredmény mögött meghúzódó rejtett késleltetések bizonyos korlátok között egy konstans erejéig eltolhatóak. Például az (X_1, X_2, X_3) és az $(X_1 + \epsilon, X_2 - \epsilon, X_3 - \epsilon)$ rejtett változók ugyanazt a mérési eredményt adják, minden értelmese ϵ -ra, amelyre a korlát: $0 < \epsilon \leq \min(X_2, X_3 + P/C)$. Ezért a probléma invertálása során csak a mért hálózati késleltetések fluktuációját tudjuk komponenseire bontani. Egészséges feltételezés, hogy minden célpont

felé legalább egy próba üres hálózaton halad keresztül és sorbanállás nélkül érkezik célba, így a hálózati késleltetés fluktuációi a 3.1. képlet értelmében az útvonal egymás után fűzött soraiban eltöltött változó mértékű várakozások kumulatív eredménye, hiszen a *propagáció*s és a kiszolgálási idők összegét állandónak tekinthetjük. Méréseink – és tőlünk független csoportok mérései [64, 65] – szerint a mai Interneten ez a feltételezés megállja helyét. Ezért a mérési idősorok konstans részét kiszervezzük a mérési adatokból:

$$Y_\nu^i \mapsto Y_\nu^i - \min_{i \in \{1, \dots, N\}} (Y_\nu^i),$$

és a transzformált adatokra az 5.1. képletet átrendezve, mátrix formalizmusba írhatjuk:

$$\begin{pmatrix} Y_2^i \\ Y_3^i \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_1^i \\ X_2^i \\ X_3^i \end{pmatrix}, \forall i \in \{1, \dots, N\}. \quad (5.2)$$

A feladat célja az 5.2. képletnek statisztikai értelemben eleget tevő együttes eloszlást, azaz a rejtett paraméterek eloszlásait megtalálni. Ennek érdekében a rejtett és a mért idősorokat kvantáljuk és az adatokat a következő szabály szerint q méretű binekbe szervezzük:

$$\begin{aligned} x_\nu^i &:= \lfloor \frac{X_\nu^i}{q} + \frac{1}{2} \rfloor, \\ y_\lambda^i &:= \lfloor \frac{Y_\lambda^i}{q} + \frac{1}{2} \rfloor, \end{aligned}$$

ahol $\lfloor \cdot \rfloor$ a lefelé kerekítés művelete, $\nu \in \mathcal{G}' := \{1, 2, 3\}$ a gráf csúcsai a gyökér nélkül, $\lambda \in \{2, 3\}$ a gráf leveleit jelöli. A kvantálás után az idősorok a $\{0, 1, 2, \dots, B\}$ egészek halmazán veszik fel értékeiket, ahol $B = \max_{\nu, \lambda, i} (y_\lambda^i, x_\nu^i)$. A kiértékelés során q értékét ügyesen úgy kell megválasztani, hogy a binekhez tartozó eseményszámok ne legyenek alacsonyak a kiértékelés ingadozásaihoz képest.

Jelölje $P_{\nu, k}$ annak a valószínűségét, hogy a ν jelű szakaszon a sorbanállási késleltetés t ideig tart:

$$P_{\nu, k} := \mathcal{P}(\mathbf{x}_\nu = k) = \mathcal{P}((k-1/2)q < t \leq (k+1/2)q).$$

A valószínűségek normáltsága miatt minden ν -re teljesül, hogy $\sum_{k=0}^B P_{\nu, k} = 1$.

A kvantálás lehetővé teszi, hogy a $P_{\nu, k}$ paraméterek segítségével egy tetszőleges mérés valószínűségét kifejezzük. A rejtett paramétereket leíró valószínűségek halmazát jelölje $\Theta := \{P_{1, k}, P_{2, k}, P_{3, k}\}_{k=0}^B$. Annak a valószínűségét, hogy egy mérés során a próbapár csomagjai a várakozásokból eredő (aq, bq) körüli késleltetést szenvednek, a következő konvolúciós képlet határozza meg:

$$\Pi_{a, b} := \mathcal{P}(\mathbf{y}_2 = a, \mathbf{y}_3 = b | \Theta) = \sum_{k \in H} P_{1, k} P_{2, (a-k)} P_{3, (b-k)}, \quad (5.3)$$

ahol H halmaz definíciójában figyelembe vesszük, hogy minden várakozási késleltetés nemnegatív ideig tart, azaz

$$H = \{0 \leq k \leq B\} \cap \{0 \leq (a - k) \leq B\} \cap \{0 \leq (b - k) \leq B\}.$$

Az 5.3. képlet felírásánál figyelembe vettük, hogy a mérések egymástól függetlenek, és feltételeztük azt is, hogy a várakozási sorok viselkedései között nincs térbeli csatolás.

Jelölje együttesen $\mathbf{Y} := \{(y_2, y_3)^{(i)}\}_{i=1}^N$ a kimért, transzformált és kvantált késleltetéseket. Az ismeretlen Θ valószínűségek meghatározásához felhasználjuk azt a tényt, hogy legjobb becslés esetén Θ maximalizálja a likelihood-függvényt, és a monotonitás miatt annak logaritmusát is. Tehát a likelihood-függvény logaritmusa:

$$\log \mathcal{L}(\mathbf{Y}|\Theta) = \sum_{i=1}^N \log \sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{y} = (y_2^i, y_3^i); \underline{x} = \underline{x}|\Theta), \quad (5.4)$$

ahol az χ^i halmaz az adott i mérés adataihoz az 5.2. egyenlet alapján a rejtett adatok összes lehetséges kombinációját tartalmazza, gyakorlatilag a jobb oldalon szereplő kifejezésben a log argumentuma az 5.3. konvolúciós képlet kiértékelése $\Pi_{y_2^i, y_3^i}$ -re.

Bár az 5.4. kifejezés szerint, a likelihood-függvény logaritmusának szélsőérték problémája analitikusan nem megoldható, léteznek numerikus algoritmusok, melyek segítségével célt érünk. Coates és Nowak munkájukban [53] az EM-algoritmus [66] használatát javasolják, mi is hasonlóan járunk el. Az iteratív eljárásként megvalósítható EM-algoritmus célja, hogy a háttéreloszlás Θ paramétereit változtatva megkeresse a likelihood-függvény maximumát. Az algoritmus minden iterációs lépése két részből áll, egy E-lépést egy M-lépés követ. Az E-lépésben (*expectation*) a paraméterek korábbi becslése, azaz az *a priori* valószínűségek ismeretében feltételes valószínűségeket képzünk a rejtett változókra, esetünkben az \mathbf{Y} szárain lévő sorbanállási időkre. Az M-lépésben (*maximization*) megkeressük a likelihood-függvény szélsőértékét az E-lépéssel kiegészített teljes adatrendszerre.

Az 5.4. kifejezés szélsőértéke ($\forall \nu \in \mathcal{G}'$ és $\forall k \in \{0, \dots, B\}$ mellett) teljesíti a

$$\begin{aligned} & \frac{\partial \log \mathcal{L}(\mathbf{Y}|\Theta)}{\partial P_{\nu,k}} \\ &= \frac{\partial}{\partial P_{\nu,k}} \left[\sum_{i=1}^N \log \sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{y} = (y_2^i, y_3^i); \underline{x} = \underline{x}|\Theta) + \sum_{\mu \in \mathcal{G}'} L_{\mu} \left(1 - \sum_{l=0}^B P_{\mu,l} \right) \right] \\ &= 0 \end{aligned}$$

feltételt, ahol a normálás feltételét L_{μ} Lagrange-multiplikátorokkal vesszük figyelembe. A képletek olvashatósága érdekében a valószínűségi változók jeleit nem

minden esetben írjuk ki. A deriválást végrehajtva

$$\begin{aligned}
0 &= \sum_{i=1}^N \frac{1}{\sum_{\underline{x} \in \chi^i} \mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta)} \frac{\partial}{\partial P_{\nu, k}} \left(\sum_{\underline{x} \in \chi^i} \mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta) \right) - L_\nu \\
&= \sum_{i=1}^N \left[\sum_{\underline{x} \in \chi^i} \left(\frac{\mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta)}{\sum_{\underline{x}' \in \chi^i} \mathcal{P}((y_2^i, y_3^i); \underline{x}' | \Theta)} \frac{\partial}{\partial P_{\nu, k}} \log \mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta) \right) \right] - L_\nu \\
&= \sum_{i=1}^N \left(\sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{x} | (y_2^i, y_3^i); \Theta) \frac{\partial}{\partial P_{\nu, k}} \log \mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta) \right) - L_\nu \quad (5.5)
\end{aligned}$$

A $\mathcal{P}((y_2^i, y_3^i); \underline{x} | \Theta)$ együttes valószínűséget Bayes-tétel szerint felírhatjuk szorzat alakban: $\mathcal{P}((y_2^i, y_3^i) | \underline{x}) \mathcal{P}(\underline{x} | \Theta)$, ahol az első tényező Θ függetlensége miatt a deriválás során az ebből származó tag eltűnik. A várakozási sorok térbeli és a minták időbeli függetlenségét felhasználva, a szorzat másik tényezője faktorizálható:

$$\mathcal{P}(\underline{x} | \Theta) = \mathcal{P}(\mathbf{x} = (x_1, x_2, x_3) | \Theta) = \prod_{\nu \in \mathcal{G}'} P_{\nu, x_\nu}. \quad (5.6)$$

Az 5.6. kifejezést visszahelyettesítve az 5.5. képletbe kapjuk, hogy

$$\begin{aligned}
0 &= \sum_{i=1}^N \left[\sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{x} | (y_2^i, y_3^i); \Theta) \frac{\delta_{k, x_\nu}}{P_{\nu, k}} \right] - L_\nu \\
&= \frac{1}{P_{\nu, k}} \left[\sum_{i=1}^N \mathcal{P}(\mathbf{x}_\nu = k | (y_2^i, y_3^i); \Theta) \right] - L_\nu.
\end{aligned}$$

A kifejezést átrendezve, és a normalitás feltételét felhasználva a Lagrange-multiplikátor értékének N adódik. Egyszerűsítés után a következő egyenletrendszەرhez jutunk, (minden ν -re és k -ra):

$$P_{\nu, k} = \frac{1}{N} \sum_{i=1}^N \mathcal{P}(\mathbf{x}_\nu = k | (y_2^i, y_3^i); \Theta), \quad (5.7)$$

ahol a $\mathcal{P}(\mathbf{x}_\nu = k | \mathbf{y} = (y_2^i, y_3^i); \Theta)$ feltételes valószínűség szemléletes jelentése az, hogy az i -ik mérésben kimért (y_2^i, y_3^i) kvantált adatpár mellett, mi annak a valószínűsége, hogy a ν élen kq ideig várakozott a mintacsomag. Újra alkalmazva Bayes-tételét, ezek a feltételes valószínűségek kifejezhetők

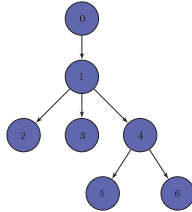
$$\begin{aligned}
\mathcal{P}(\mathbf{x}_1 = k | \mathbf{y} = (a, b); \Theta) &= \frac{P_{1, k} P_{2, a-k} P_{3, b-k}}{\Pi_{a, b}}, \\
\mathcal{P}(\mathbf{x}_2 = k | \mathbf{y} = (a, b); \Theta) &= \frac{P_{1, a-k} P_{2, k} P_{3, b-a+k}}{\Pi_{a, b}}, \\
\mathcal{P}(\mathbf{x}_3 = k | \mathbf{y} = (a, b); \Theta) &= \frac{P_{1, b-k} P_{2, a-b+k} P_{3, k}}{\Pi_{a, b}}, \quad (5.8)
\end{aligned}$$

alakban. Az 5.3., az 5.8. és az 5.7. kifejezések csatolt egyenletrendszert alkotnak, az EM-algoritmus ezeknek az egyenleteknek az iteratív kiértékelését jelenti. A numerikus kiértékelést akkor állítjuk le, ha Θ paraméterek rendben konvergáltak. A kilépési feltételt úgy rójuk ki, ha egy iteráció során a legnagyobb relatív paraméterváltozás is alulmarad egy előre definiált küszöbértéknél, akkor megtaláltuk a megoldást. Vesszőzve az előző iteráció megfelelő paramétereit, a megállási feltétel relációja

$$\max_{\nu \in G', k \in \{0, \dots, B\}} \left| \frac{P_{\nu, k} - P'_{\nu, k}}{P'_{\nu, k}} \right| B \leq \varepsilon.$$

5.2.1.1. Tetszőleges fa lefedése Y-okkal

Ebben az alfejezetben bemutatjuk, hogy tetszőlegesen nagy feszítőfa éleit jellemző sorbanállási késleltetés megbecsülhető az Y-módszer kiterjesztésével. Ket-tőnél több levél esetén a vizsgálat tárgyát képező feszítőfa szegmenseinek egy részét leíró valószínűségek közvetlenül megkaphatóak az Y-ra kidolgozott módszer eredményeiből, míg a többi él paraméterei az átfedő szakaszok valószínűség-sűrűségeinek megfelelő numerikus dekonvolúciójával határozhatóak meg. A felbonthatóság természetes feltétele, hogy a csomagpár mérések közül legalább egy mérés érintse az összes végpontot illetve az összes elágazást. A konkrét mérési megvalósításnál nem törekedtünk a redundanciára, minden lehetséges párra elvégeztük a méréseket. Ebből következően a teljes lefedés megvalósul, sőt számos szegmensre több, mint egy becslési eredményt is kaptunk. Ahol egy szegmensre több becslés áll rendelkezésünkre, az adatokat a becslési módszer hibájának meghatározására is felhasználjuk.



5.3. ábra. Egy lehetséges példa az egy forrás, négy nyelő mérési elrendezésre. A levél párok tetszőleges kombinációját megcímezve küldjük a hálózatba a próbákat a forgalom jellemzőit letapogatni. Az élek egy részhalmazára (például $0 \rightarrow 1$, $1 \rightarrow 2$, ...) az Y-esetre kifejlesztett EM-algoritmust alkalmazva megkapjuk a sorbanállási idők valószínűségeit. Majd a közvetlenül nem felbontható szakasz paraméterei numerikus dekonvolúcióval meghatározhatóak, így a $1 \rightarrow 4$ kifejezhető az $1 \rightarrow 5$ és az $4 \rightarrow 5$ paramétereinek segítségével.

Az 5.3. ábra példáját elemezve elegendő elvégezni a (2, 3) és (5, 6) levelekre az Y-os csomagpár mérést, természetesen felváltva, hogy biztosítani tudjuk az $0 \rightarrow 1$ szakasz állapotát a két mérőfolyam „egy időben” tapogatja le. Az Y-esetre kidolgozott módszerrel közvetlen becslést kapunk a $P_{1,k}, P_{2,k}, P_{3,k}, P_{5,k}, P_{6,k}$ valamint $P'_{4,k}$ eloszlásokra, ahol a $P'_{4,k}$ eloszlás a (5, 6) mérés adataiból feltáráshoz $0 \rightarrow 4$ összetett szakaszt jellemzi, ami ebben az esetben két valószínűségi változó összegének eloszlása, azaz:

$$P'_{4,k} = \sum_{\kappa \in \{0, \dots, k\}} P_{1,k-\kappa} P_{4,\kappa}.$$

A kifejezésben $P_{4,\kappa}$ változók a még ismeretlen keresett mennyiségek. A numerikus dekonvolúciót a nemnegatív legkisebb négyzetek módszerének [67] segítségével határozzuk meg.

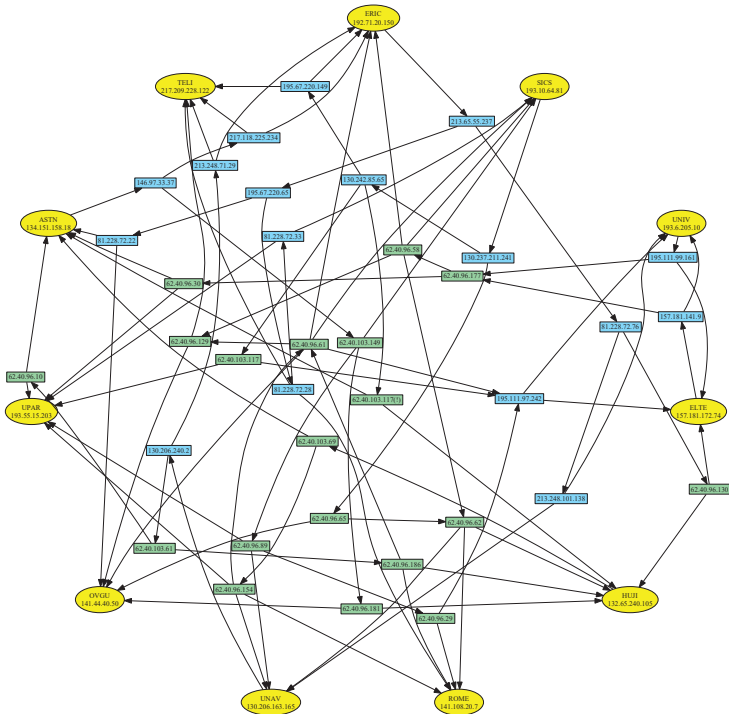
5.2.2. Mérési eredményeink

A becslési eljárás teljesíthetőségét és konvergencia tulajdonságait először különböző eloszlású véletlen számokon, *ns* hálózati szimulátoros adatokon illetve laboratóriumi hálózatban lefolytatott méréseken vizsgáltuk. Kutatásaink az elméleti és a becslt eloszlások között jó egyezést állapítottak meg, az eloszlás átlagára megbecsült hiba a választott q binelési méret alatt volt.

Első sikeres nagyskálájú tomográfias méréseinkben 11 mérőpontot¹ használtunk a 3.4. fejezetben bemutatott Etomic mérőhálózatból. A mérőpontok között kifeszülő logikai hálózat 36 elágazási pontot és 92 összekötő élt tartalmazott, amit az 5.4. ábrán mutatunk be. Az elágazási pontok közül 19 az Európai Akadémiai Hálózatban, a *GÉANT* [68] hálózatban üzemelt, melyre a gyors (2–15 *Gbps*) hálózati kiszolgálási sebesség jellemző. A 7 forrás² Etomic mérőpont egyidőben a 3.3.2. fejezetben bemutatott módszerrel használva szorosan illeszkedő kisméretű, $-P = 40$ *byte*-os, UDP protokollú, – csomagpárokat szórt a párokat alkotó címeket folyamatosan keverve, ügyelve arra, hogy a párok minden lehetséges kombinációja előforduljon. A csomagpárok között 10 *ms* várakozási időt iktattunk be. A mintacsomagok indulási és érkezési időbélyegei alapján megszerkesztettük a hálózati késleltetések csatolt idősorait. Azokat az adatokat, ahol az egyik idő-adat hiányzott, akár hálózati veszteség, akár program indítási csúszások miatt, nem vettük figyelembe. A mérés nyers idősorából a kis számban előforduló kilógó pontokat mérési hibának elkönyvelve kiszűrtük. Az outlier-szűrés mérésenként kevés (< 10) pontot érint. Eredményeink lényegében nem módosulnának, ha a kiértékelésben hagynánk ezeket a pontokat, ám azokat az adatok között hagyva jelentősen megnövekedne B értéke, ami a numerikus kiértékelés futásidején sokat rontana. A

¹A mérésben részt vevő mérőpontok listája: *fericsson, sics, univet, elte, jerusalem, rome, pamplona, magdeburg, paris, birmingham, telia*.etomic.org

²A bemutatott mérésben 4 végpont csak nyelőként üzemelt.



5.4. ábra. Az Interneten levezényelt nagyskálájú Y-tomografikus mérés logikai kapcsolata. A sárga ovális csúcsok a mérőpontokat jelölik. A *GEANT* akadémiai hálózat útvonalválasztó eszközeit az ábrán zöld színű téglalappal szedtük, a többi elágazási pontot kékkel. A nyilak a mérőcsomagok közlekedési irányát szemléltetik.

mérésenként fennmaradó több mint 8000 elemet számláló idősorokat kiértékeltek a bemutatott Y-tomográfias módszerrel. A becslés eredményeként a vizsgált hálózat belsejére jellemző sorbanállási idők eloszlását becsültük meg. A becslés során először önkényes kezdeti értéket adunk Θ minden paraméterének. Az *a priori* valószínűségeket egyenletes valószínűségnek állítottuk be, azaz $P_{\nu,k} = 1/(B+1)$ minden ν -re és k -ra.

Amint azt az eljárás kibővítésénél említettük, léteznek olyan szegmensek, amik

az Y-ok különböző kombinációjában is előfordulnak, így ezekre több becsült sorhossz eloszlást kapunk eredményül. Ezt egyrészt arra használjuk, hogy megbizonyosodjunk a becslési módszer konzisztensségéről, másodsorban ebből számoljuk a becslés hibáját. A többször lefedett szegmensekre az átlagos eloszlást tekintjük végeredményként.

5.2.2.1. Analízis

Az adatok együttes ábrázolásának érdekében az eloszlások első két momentumát vezetjük fel a feszítőhálózat szegmenseire.

Az 5.5. és az 5.6. ábrákon található adatok érdekes belső szerkezetet mutatnak. Első észrevételünk, hogy az állapotot leíró rejtett paramétersokaság mintegy két nagyságrendet ölel fel. Az átlagos sorbanállási idő $\approx 5 \mu s$ értékétől az 1 ms nagyságú várakozási időig húzódik. A gyors (több gigabites) Akadémiai Hálózat gerincvonalán elhelyezkedő sorokra jellemzőek a kisebb, míg a bekötő vonalakon hosszabb várakozások fordulnak elő.

Érdekes geográfiai megfigyelés, hogy a legnagyobb ingadozással terhelt várakozási sorok a déli régiókban, például a jeruzsálemi és a pamplonai egyetem bekötő vonalainál tapasztalhatóak, itt alakulnak ki hálózati torlódások. Megfigyeléseinkből látszik a hálózati vonalak aszimmetriája is, a mérőpontokra befutó vonalak várakozási sorai nagyobbak, mint a kimenő vonalaké. Az erre adható egyik lehetséges magyarázat például, hogy azokban az intézményekben, ahol az Etoomic mérőpontok üzemelnek, az Interneten keresztül letöltött információ mennyisége meghaladja a saját kiszolgálású adatok mennyiségét. A vonalakra becsült adatok térbeli struktúráját tanulmányozva látszik, hogy a hálózat magját alkotó, a belső elágazási pontokat összekötő vonalakon a legkisebb a sorbanállási idő várható értéke és szórása, ami nem meglepő, hiszen a gerinchálózat a leggyorsabb vonalakból és eszközökből épülnek fel.

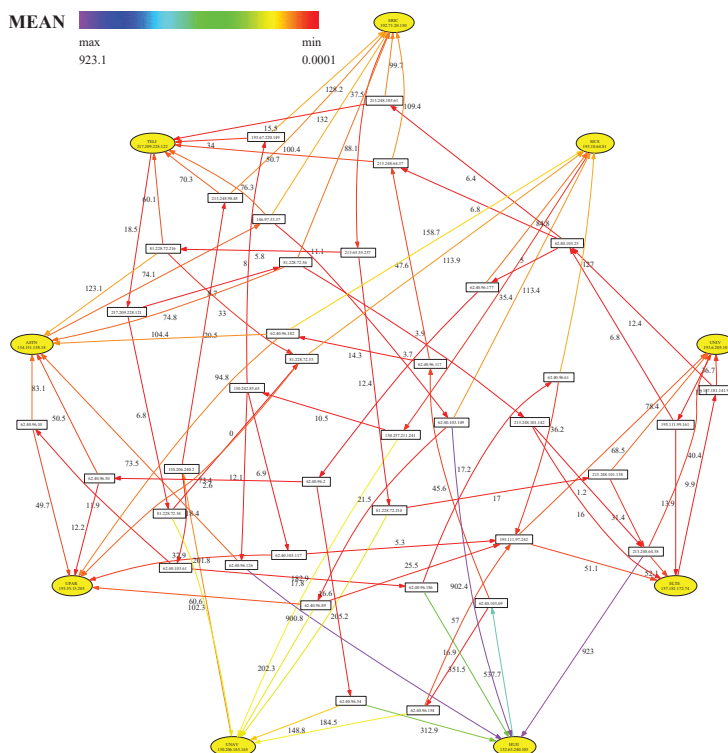
A sorbanállási idők várható értékei ($E[x]$) a log-normális eloszlást követik,

$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(\ln x - m)^2}{2\sigma^2}}$, így a kiegészítő kumulatív eloszlásra jól illeszkedik a

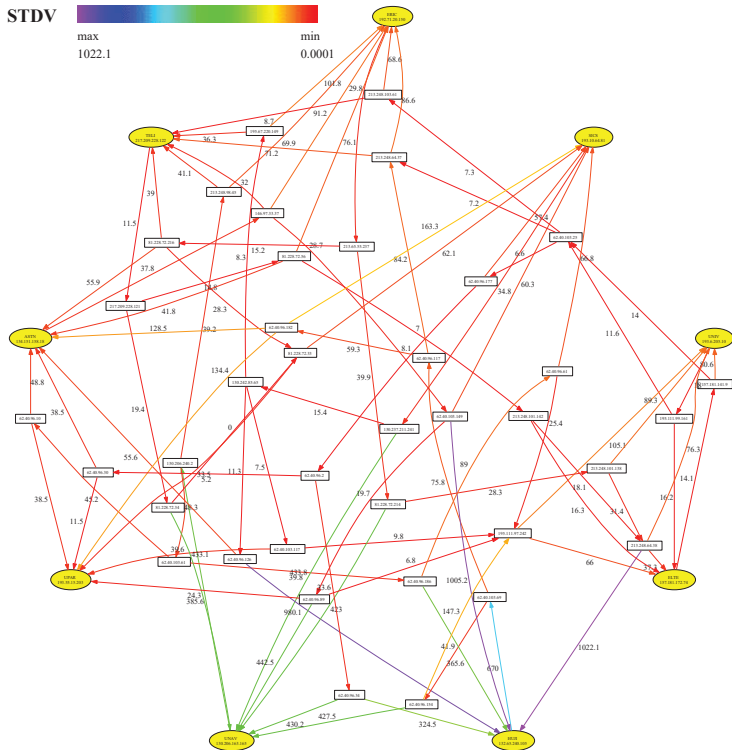
$$P(\mathbf{x} < x) = \frac{1}{2} \left[1 - \operatorname{erf} \left(\frac{\ln x - m}{\sigma\sqrt{2}} \right) \right], \quad (5.9)$$

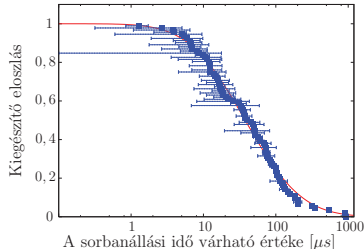
kifejezés, amelynek illesztési paramétereire $\sigma \approx 1.51 \mu s$ és $m \approx \ln(57.8 \mu s)$ adódik, ahogy az 5.7. ábrán látható. A korábban megállapított térbeli tagolódás ellenére az egybe boronált adatok síma folytonos függvényt követnek, anélkül, hogy a pontok az 5.9. görbe különböző részein csoportosulnának, ritkább-sűrűbb részeket kialakítva.

Az 5.8. ábrán a sorokat jellemző szórás és várható értékeket ábrázoltuk egymás függvényében. A diagonális közelében sűrűsödő pontok arra utalnak, hogy a ponthoz tartozó él sorhosszát exponenciális illetve közel exponenciális eloszlású valószínűségek jellemzik. Azok a pontok, melyek az átlótól távolabb helyezkednek

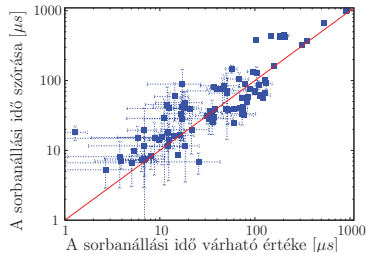


5.5. ábra. Az Y-tomográfiás becsléssel meghatározott sorbanállási idők várható értéke. Az éleken és a skálám feltüntetett időadatok mértékegysége μs .





5.7. ábra. A becslt sorbanállási idők várható értékeiből készített kiegészítő kumulatív eloszlás. A hibát az egy szegmenshez tartozó több kiértékelt adat szórásával definiáltuk. A folytonos vonallal a 5.9. képlettel illesztett görbét ábrázoljuk.



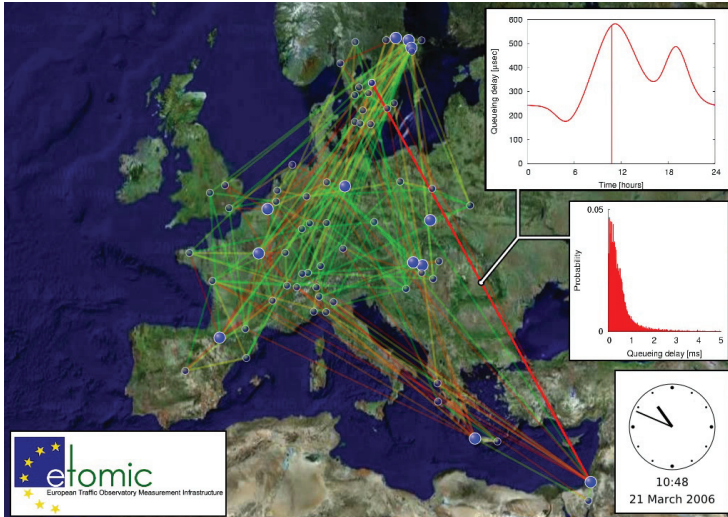
5.8. ábra. A sorbanállási idők szórása a várható értékek függvényében. A folytonos vonal a diagonálist jelöli.

el, jobban jellemezhetőek hatványfüggvény szerinti lecsengéssel, ami önhasonló természetű forgalomra utal.

5.2.2.2. Az Internet forgalom egy napja

A dolgozat mellékletét képezi egy animációs film, mellyel az európai Internet egy teljes napját (2006. március 21-t) felölő forgalmi viselkedést mutatjuk be. A mérésorozatot az Etomic mérőhálózattal a nap folyamán egyenletesen elosztva, 3 óránként végeztük.

Méréseink és az 5.2. alfejezetben bemutatott Y-tomográfias becslési eljárás segítségével feltérképeztük és rekonstruáltuk az Interneten a várakozási sorok késleltetéseit. Az animációs filmből egy kockát az 5.9. ábrán mutatunk be. A



5.9. ábra. Torlódásmintázat az Interneten (részlet az animációs filmből).

mérésben részt vevő végpontokat nagyobb méretű, lila korongokkal jelöljük, az elágazási pontokat kisebb méretű, kék korongokkal ábrázoljuk. A mérések napján a hálózat topológiájában nem találtunk a feszítőhálózat struktúráját érintő lényegi változást. Így lehetőségünk nyílt arra, hogy statikus szerkezetű hálózatot ábrázoljunk a mozgógépeken, azaz minden gráf él teljes napi viselkedését nyomon követhetjük, anélkül, hogy új élek jelennének meg vagy egyes szegmensek eltűnnének az ábrázolás során. A gráfok csúcsait alkotó eszközöket földrajzilag helyesen vetítettük rá Európa térképére, és a gráf éleit a megbecsült sorbanállási idő várható értékei alapján színeztük. A mérések közötti időpillanatokra az élek színezését interpoláció segítségével határoztuk meg. Napi ismétlődést feltételezve minden élhez rendelt idősort periodikus *spline*-nal illesztettünk, és azt az animációs kockák időpontjaiban mintáztuk. Vörös árnyalattal színeztük azokat a vonalakat, ahol nagy a sorbanállási idő várható értéke, és zölddel a dugulatlan sorokat.

Az Internet torlódás szempontjából megvizsgált állapotát az animációs filmen vetítjük le egymás után, először a teljes összképet bemutatva, majd két önkényesen választott élet ragadunk ki. Elsőként egy, a hálózat peremét a gyors maghálózattal összekötő szegmenst választunk, másodikként két belső elágazási pontot

összekötő belső vonalat ragadunk ki. A kiválasztott vonalak színezése mellett feltüntetjük a becslésekből származó $P_{v,k}$ sűrűségfüggvényeket és nyomon követjük a *spline*-nal illesztett átlag várakozási idő napi menetét.

Az egyedi szegmensek dugult állapotának megfelelő hosszabb várakozási sorokat két ízben tapasztalunk a nap folyamán. Az sorhosszak átlagának csúcsai 12 és 19 óra környékén találhatók a közép európai időzónában. A jelenséget magyarázhatjuk az emberi tevékenység és a hálózat kölcsönhatásának eredményeként. Valószínűsítve, hogy a dél környéki pihenő időben egyszerre sok ember rutinszerűen elolvassa a hálózaton az aznapi híreket, elvégzi magánlevelezését, stb. magyarázhatja a 12-es csúcsot. A 19 órai csúcs mögött az érvelés hasonló, ezidőtájt érnek tipikusan haza a munkából az emberek, s otthonról használják intenzíven a hálózatot.

5.3. Tomográfiás sorhosszeloszlás-becslés fán

A sorbanállási idő valószínűségeiről számot adó, az 5.2. fejezetben bemutatott becslési módszer segítségével beláttuk, hogy feltárhatjuk a mérő végpontok között kifeszülő hálózat belső szegmenseire a forgalom dinamikáját jellemző statisztikai jellemzőket. A mérési módszer skálázás szempontjából problémás, növelve a mérőpontok számát, exponenciálisan megnövekedik a szükséges minták száma, megnövelve a hálózat terhelését és a mérésre fordított időt. Amennyiben a mérés kezdetén jól ismerjük a hálózat topológiáját, kevesebb mérőpár elegendő a mérés elvégzésére, de ezek könyvelése és ellenőrzése nem egyszerű feladat, a mérés levezénylésénél igényelne kombinatorikai kiértékelést.

Ebben a szakaszban azt vizsgáltuk meg, hogy az Y-kiértékelés likelihood-függvényen alapuló megvalósítása, hogyan általánosítható tetszőleges mérőfára. Az 5.4. képlet kiterjeszthető tetszőleges hálózatra, de a második összegzésnél (χ^i elemekre) a rejtett változók összes lehetséges értékére el kell végezni a kiértékelést, ami exponenciálisan növekvő bonyolultságú probléma. Liang és Yu [69] munkájukban pseudo-likelihood statisztikát vezetnek be, és a becslési eljárásuk során ezt a statisztikát maximalizálják. Cikkükben bemutatják, hogy a bevezetett rekurzív statisztika, akárcsak a maximum likelihood alapú statisztika, konzisztens és aszimptotikusan normális becslést ad. Megmutatjuk, hogy a likelihood alapú statisztikát megfelelő matematikai átalakításokkal olyan az EM-algoritmusra alakíthatjuk, melynek iteratív kiértékelése polinomiális számításigényt támaszt.

Az Y-tomográfiás méréshez képest technikailag most nem csomagpárokat kell küldeni a hálózatba, hanem hosszabb, szorosan illeszkedő csomagokból felépített vonatokat. A hálózati késleltetést mérő program, amit a 3.3.2 alfejezetben mutatunk be, képes ezt a feladatot ellátni. A program jósági tesztje során megtanultuk, hogy 18–20 csomag hosszúságú vonatot még kiváló korrelációs tulajdonságokkal küld ki a hálózatba. Az Etomic mérőrendszerben jelenleg ennyi végpont üzemel, így a mérőprogramot módosítás nélkül használhatjuk arra, hogy egy forrásból

Jelölés	A jelölés értelmezése
$\nu \in \mathcal{G}$	a fa egy tetszőleges csúcsa, $\nu = 0$ a fa gyökere
π^ν	a $\nu \neq 0$ csúcs őse
Γ^ν	a csúcs gyermekeinek halmaza: $\{\mu \in \mathcal{G} : \pi^\mu = \nu\}$
Λ^ν	a ν csúcsból elérhető levelek halmaza
Λ^0	a fa leveleinek halmaza
$\Lambda^{\nu \setminus \mu}$	levelek, amik ν -ből elérhetők, de μ -ból nem: $\Lambda^\mu \setminus \Lambda^\nu$
\mathbf{d}_ν	valószínűségi változó: a $0 \rightarrow \nu$ szakasz sorban állási ideje
y_λ^i	a $\lambda \in \Lambda^0$ levélnél i -ként mért teljes sorban állási idő
$\underline{\mathbf{y}} \stackrel{A}{=} \underline{\mathbf{y}}^i$	a mérés adatainak leszűkítése, azaz $\forall \lambda \in A \subset \Lambda^0: \mathbf{y}_\lambda = y_\lambda^i$
Θ	a rejtett paraméterek halmaza $\{P_{\nu,k} : \nu \in \mathcal{G} \setminus \{0\}; 0 \leq k \leq B\}$

5.1. táblázat. A *message passing* módszer bemutatásánál használt jelölések gyűjteménye.

egyszerre az összes többi végpontra mérőforgalmat küldjünk, egy méréssel letagatva a teljes feszítő fát.

A következő alfejezetben a fákra kidolgozott EM-algoritmusra épülő becslési módszert mutatjuk be. A tomográfiás-becslés fákra kiterjesztett modelljében a következő feltételezésekkel élünk:

- A mérésből származó hálózati késleltetések valószínűségi változói ($\underline{\mathbf{y}}$) függetlenek minden egyes i mérési időpontra,
- A mérési kísérletek időtartama alatt a hálózat topológiája (\mathcal{A}) nem változik, és ezt a lehetőségekhez mérten méréseinkkel ellenőrizzük,
- Minden belső szegmens sorát jellemző késleltetés (\mathbf{x}_ν) független azonos eloszlásból ($P_{\nu,k}$) húzott valószínűségi változó, melyeket együttesen Θ jelölünk.

5.3.1. A becslési módszer

A becslési módszer EM-algoritmusának levezetésénél a probléma gráf jelleget használjuk fel az 5.4. kifejezés általánosított alakjának egyszerűsítésére. Az átalakítások során felmerülő jelöléseinket az 5.1. táblázatban foglaljuk össze a könnyebb visszakereshetőség kedvéért. A forrás és a fa leveleinél található végpontok között lemért hálózati késleltetéseket $\underline{\mathbf{y}}^i$ -vel jelöljük, míg $\underline{\mathbf{x}}^i$ azokat a belső késleltetéseket jelenti, amiket közvetlenül nem ismerhetünk. Mindkét jelölésnél $i = 1 \dots N$ a mérési minta sorszámát jelenti.

A becslés során a likelihood-függvényt maximalizáljuk. A független mérések

feltételéből adódóan a likelihood-függvény általánosan

$$\mathcal{L}(\mathbf{Y}|\Theta) = \prod_{i=1}^N \mathcal{P}(\underline{y} = \underline{y}^i|\Theta),$$

szorzat alakjában írható, ahol a hálózat topológiájából eredő minden konvolúciós tudás a

$$\mathcal{P}(\underline{y} = \underline{y}^i|\Theta) = \sum_{\underline{x}} \mathcal{P}(\underline{x} = \underline{x}|\Theta) \mathbf{1}_{\underline{y}^i = \mathcal{A}\underline{x}^i} = \prod_{i=1}^N \sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{y} = \underline{y}^i; \underline{x} = \underline{x}|\Theta), \quad (5.10)$$

χ^i halmazba van belekódolva, amit az $\mathbf{1}_{\underline{y}^i = \mathcal{A}\underline{x}^i}$ indikátor ír elő az 5.2. képlet analógiájára.

Az 5.5. levezetési lépéseit végrehajtva a fákra általánosított esetben a

$$\sum_{i=1}^N \left(\sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{x}|\underline{y}; \Theta) \frac{\partial}{\partial P_{\nu,k}} \log \mathcal{P}(\underline{y}; \underline{x}|\Theta) \right) - L_\nu = 0$$

feltételt kapjuk a likelihood-függvény szélsőértékére.

A $\mathcal{P}(\underline{y}; \underline{x}|\Theta)$ valószínűség meghatározásánál Bayes-tételt alkalmazva, majd a mintavételezés és a belső sorok függetlenségét kihasználva szorzat alakba írhatjuk, úgy mint az 5.6. képletnél, és a

$$\begin{aligned} 0 &= \sum_{i=1}^N \left[\sum_{\underline{x} \in \chi^i} \mathcal{P}(\underline{x} = \underline{x}|\underline{y} = \underline{y}^i; \Theta) \frac{\delta_{k,x_\nu}}{P_{\nu,k}} \right] - L_\nu \\ &= \frac{1}{P_{\nu,k}} \left[\sum_{i=1}^N \mathcal{P}(\mathbf{x}_\nu = k|\underline{y} = \underline{y}^i; \Theta) \right] - L_\nu, \end{aligned}$$

általános feltételrendszerre jutunk, amit $P_{\nu,k}$ -ra átrendezve kapjuk, hogy

$$\begin{aligned} P_{\nu,k} &= \frac{1}{L_\nu} \sum_{i=1}^N \mathcal{P}(\mathbf{x}_\nu = k|\underline{y} = \underline{y}^i; \Theta) \\ &= \frac{1}{L_\nu} \sum_{i=1}^N \frac{\mathcal{P}(\underline{y} = \underline{y}^i; \underline{x}_\nu = k|\Theta)}{\mathcal{P}(\underline{y} = \underline{y}^i|\Theta)}. \end{aligned} \quad (5.11)$$

Az 5.11. hánadosaiban szereplő valószínűségek bonyolult konvolúciós kifejezések, minden fára másként kell képezni, az 5.3. konvolúciójához hasonlóan. Ahelyett, hogy a topológia minden egyedi megvalósulásaira különböző zárt kifejezést vezetnénk le, kihasználjuk a geometria nyújtotta adottságokat. A képlet numerikus kiértékelésénél a valószínűségekre helyi feltételeket rovunk ki és a fa

teljes bejárásával kiértékelhetővé tesszük az 5.11. törtet. Ez a kiértékelési módszer a *message passing* [70] algoritmusok családjába tartozik.

A nevezőben található $\mathcal{P}(\mathbf{y} = \underline{y}^i | \Theta)$ feltételes valószínűséget értékeljük ki az iteráció első lépéseiben, miközben a fát a leveleitől a gyökeréig felfelé járjuk be. A következőként bevezetett valószínűségeloszlás feltételrendszerén keresztül helyesen veszi figyelembe a konvolúciós szabályokat:

$$\mathcal{P}(\underline{y} \stackrel{\Lambda^\nu}{=} \underline{y}^i | \mathbf{d}_\nu = d; \Theta) = \begin{cases} \delta_{y_\nu^i, d}, & \text{ha } \nu \in \Lambda^0, \\ \prod_{\mu \in \Gamma^\nu} \sum_k P_{\mu, k} \mathcal{P}(\underline{y} \stackrel{\Lambda^\mu}{=} \underline{y}^i | \mathbf{d}_\mu = d + k; \Theta), & \text{ha } \nu \notin \Lambda^0. \end{cases} \quad (5.12)$$

Az 5.12. képlet szemléletes jelentése a következő. A fa egy tetszőleges ν pontjában elhelyezkedve megmondjuk a pontból elérhető levelekhez tartozó mérésnek bekövetkezésének valószínűségét, a minta e csúsig tapasztalt késleltetésének fényében. Természetesen a leveleken ($\nu \in \Lambda^0$) a késleltetés meg kell egyezzen a mért késleltetéssel, ezt a feltételt jelöli ki a Kronecker-delta. A fa belső ν csúcán elhelyezkedve, az 5.12. feltételes valószínűség megegyezik azzal, hogy a vizsgált csúcs gyermekeiig a minta késleltetése a ν csúcs utáni szegmensben lévő sorbanállás idejével nőtt meg.

Összehasonlítva az 5.12. és az 5.11. képlet nevezőjét, észrevesszük, hogy

$$\mathcal{P}(\underline{y} = \underline{y}^i | \Theta) = \mathcal{P}(\underline{y} \stackrel{\Lambda^0}{=} \underline{y}^i | \Theta) = \mathcal{P}(\underline{y} \stackrel{\Lambda^0}{=} \underline{y}^i | \mathbf{d}_0 = 0; \Theta).$$

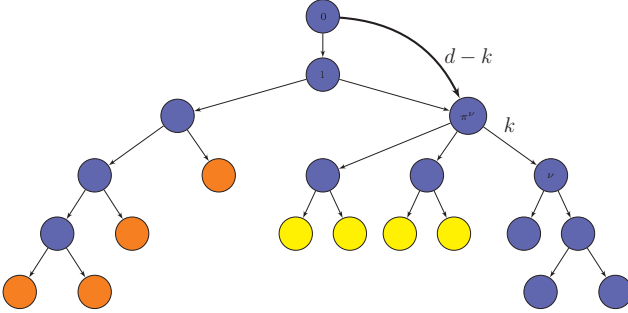
Az 5.11. képlet számlálójának meghatározásához újabb együttes valószínűségeloszlásokat kell bevezetnünk, melyeket a gráfon fordított irányban, azaz a fát fentről lefelé bejárva tudunk numerikusan kiértékelni:

$$\mathcal{P}(\underline{y} \stackrel{\Lambda^{0 \setminus \nu}}{=} \underline{y}^i; \mathbf{d}_\nu = d | \Theta) = \begin{cases} \delta_{d, 0} & \text{ha } \nu = 0, \\ \sum_k P_{\nu, k} \mathcal{P}(\underline{y} \stackrel{\Lambda^{0 \setminus \nu}}{=} \underline{y}^i; \mathbf{d}_{\pi^\nu} = d - k | \Theta) & \text{ha } \nu \neq 0. \end{cases} \quad (5.13)$$

Az 5.13. képlet egy tetszőleges ν csúcsig tartó késleltetés valószínűsége és az e pontból nem elérhető levelek mérési megfigyelései között teremt kapcsolatot. Természetesen a fa gyökeréből nem elérhető levelek üres halmazt alkotnak, ezért $\nu = 0$ csúcsban az 5.13. valószínűség értéke 1, ha $d = 0$. E feltétel természetes jelentése, hogy a mintáknak indulásuk pillanatában semmilyen késleltetése nincs. És ugyanezt az érvelést követve a komplementer eseményre a valószínűség értéke 0 minden egyéb d késleltetésre.

Az 5.13. összegben szereplő tagokban egy új együttes valószínűséget is bevezetünk, ami kapcsolatot teremt egy gráf csúcsa és ősenek késleltetési tulajdonságai között:

$$\mathcal{P}(\underline{y} \stackrel{\Lambda^{0 \setminus \nu}}{=} \underline{y}^i; \mathbf{d}_{\pi^\nu} = d - k | \Theta) =$$



5.10. ábra. A $\mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{0 \setminus \nu}} = \underline{\mathbf{y}}^i; \mathbf{d}_{\pi^\nu} = d - k | \Theta)$ valószínűség 5.14. szerinti kifejtéséhez készített szemléltetés.

$$\mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{0 \setminus \pi^\nu}} = \underline{\mathbf{y}}^i; \mathbf{d}_{\pi^\nu} = d - k | \Theta) \cdot \mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{\pi^\nu \setminus \nu}} = \underline{\mathbf{y}}^i | \mathbf{d}_{\pi^\nu} = d - k; \Theta). \quad (5.14)$$

A képlet jelentésének értelmezéséhez az 5.10. ábrát hívjuk segítségül. A ν csúcs előtti vonalon, az 5.13. képlet jobb oldala szerint, minden lehetséges k idejű várakozást számba veszünk, ami megköveteli, hogy ν őseit a minta útja $d - k$ ideig tartson. A rejtett változók ezen feltételével egyidőben megvizsgáljuk, mi annak a valószínűsége, hogy a ν csúcsból elérhetetlen leveleken az i -ik mérésben kimért adatokat látjuk. Az ábrán ezeknek a leveleknek a halmazát a sárga és narancs színnel jelölt végpontok együttese szemlélteti. Ugyanezt a valószínűség kapjuk, ha egyszerre teljesül az, hogy π^ν -ig a késleltetés $d - k$ és az innen elérhetetlen (narancsral jelölt) pontokon stimmelnek a mérési adataink, valamint π^ν -ig a minta $d - k$ nagyságú késleltetését feltételezve, a π^ν csúcsból elérhető, de ν -ból nem elérhető (citrommal jelölt) végpontok is a mért eredményeket látjuk rendben.

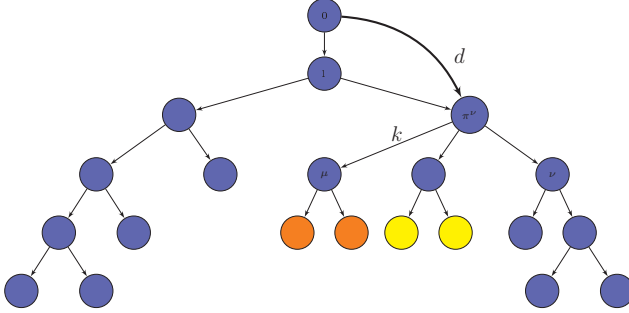
Az 5.14. szorzat második tényezője átírható

$$\mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{\pi^\nu \setminus \nu}} = \underline{\mathbf{y}} | \mathbf{d}_{\pi^\nu} = d; \Theta) = \begin{cases} \delta_{d,0}, & \text{ha } \nu \in \Gamma^0 \\ \prod_{\mu \in \Gamma^{\pi^\nu \setminus \{\nu\}}} \sum_k P_{\mu,k} \mathcal{P}(\underline{\mathbf{y}}^{\Lambda^\mu} = \underline{\mathbf{y}}^i | \mathbf{d}_\mu = d + k; \Theta). \end{cases} \quad (5.15)$$

alakba. Az 5.11. ábra szemléltet, a $\Lambda^{\pi^\nu \setminus \nu}$ halmazba tartozó végpontokat π^ν gyermekei szerint történő kifejtést.

A fent bevezetett 5.13, 5.14, 5.15. és 5.12. valószínűségek segítségével felírhatjuk az 5.11. kifejezés számlálóját is, ezzel rekurzív formulát kapunk a rejtett paraméterek, azaz az éleken található sorok hosszának valószínűségeire:

$$\mathcal{P}(\underline{\mathbf{y}} = \underline{\mathbf{y}}^i; \mathbf{x}_\nu = k | \Theta) =$$



5.11. ábra. A $\mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{\pi^\nu \setminus \nu}} \mid \underline{\mathbf{y}} \mid \mathbf{d}_{\pi^\nu} = d; \Theta)$ valószínűség 5.15. szerinti kifejtéséhez készített szemléltetés.

$$\mathcal{P}(\underline{\mathbf{y}}^{\Lambda^0} \mid \underline{\mathbf{y}}^i; \mathbf{x}_\nu = k \mid \Theta) = P_{\nu,k} \sum_d \mathcal{P}(\underline{\mathbf{y}}^{\Lambda^{0 \setminus \nu}} \mid \underline{\mathbf{y}}^i; \mathbf{d}_{\pi^\nu} = d - k \mid \Theta) \mathcal{P}(\underline{\mathbf{y}}^{\Lambda^\nu} \mid \underline{\mathbf{y}} \mid \mathbf{d}_\nu = d).$$

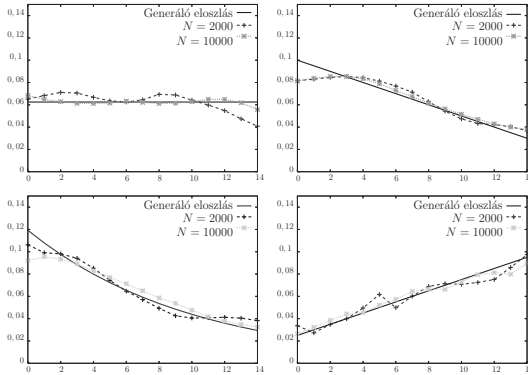
Az 5.11. rekurziós képletben nem határoztuk meg a Lagrange-multiplikátorok értékére igaz képletet, hanem utólag minden $P_{\nu,k}$ eloszlást visszanyomunk, ami ezzel teljesen ekvivalens megoldás.

Természetesen a fákra kidolgozott általánosított módszer legegyszerűbb esete az Y-alakú mérőhálózat. Ebben a határesetben Az 5.11. rekurziós formula egyszerűsítések után megnyugtatóan visszavezethető az 5.8. kifejezésre.

A levezetés során hallgatólagosan feltételeztük, hogy a gyökér fokszáma 1, amit mérési tapasztalataink is megerősítenek. Ha ez nem teljesülne, akkor szét kellene bontani az eredeti fát fokszámnyi fára, ami újra eleget tesz ennek a feltételnek, de nem szabad figyelmen kívül hagyni hogy a különválasztott fák idősorai dekorreláltak.

5.3.2. Mérési eredményeink

A fákra kidolgozott tomografikus becslés eljárását először szimulált kísérleti adatokra alkalmaztuk. Olyan fát szerkesztettünk, melynek mind 4 szintjén egy-egy csúcsból 4 kimenő vonal indul, azaz a fát összesen 85 él alkotja. Az élekhez különböző típusú és paraméterű, előre beállított eloszlásokat rendelünk, mindegyiket 15 binre osztjuk, azaz a rendszert 1275 paraméter írja le. A beállított eloszlásoknak megfelelően véletlen számokat generálunk és a vonalak mentén felösszegezve $N = 2000$ illetve $N = 10000$ szimulált mérési adatot képzünk. Az 5.12. ábrán 4 tetszőleges vonal becslült eloszlásait vetjük egybe az eredeti paraméterekkel.



5.12. ábra. Az algoritmus jószágvizsgálata szimulált adatokkal. A 85 élből álló fa 4 tetszőleges vonalának eredeti eloszlását vetjük egybe a tomográfiás kiértékelés adataival, $N = 2000$ és $N = 10000$ mérési mintára alkalmazva azt. Az ábrákon a vízszintes tengelyen a késleltetési idő k , a függőleges tengelyen annak valószínűsége szerepel.

A fákra alkalmazott tomográfiás mérésben szorosan csatolt csomagokat küldünk egy-egy forrásból a többi nyelőként üzemelő mérőpont felé, miközben mérjük minden csomag hálózati késleltetését. Az *UDP* protokollú mintacsomagok méretét ebben a mérési sorozatban is a lehető legkisebbnek állítjuk be, hogy a mérésből eredő hálózati terhelést minimalizáljuk. 48 *byte*-os csomagokat használunk, és minden mérésben $N = 10000$ vonatot indítunk útjára, melyből a minták kevesebb, mint 0,5 %-a veszik el hálózati csomagvesztés illetve a program aszinkron indulása következtében. A csomagvonatok között 10 *ms* közöket iktatunk be, hogy teljesüljön a független mintavételezésre kirótt követelmény, és relaxálhasson a hálózat. Az elmúlt mintegy három évre visszamenőleg, fás méréseink száma meghaladja a 2000 mérést, melyek nyersadatait és a kiértékelte adatok nagy részét a 4. fejezetben bemutatott virtuális obszervatórium adatbázisaiban is tároljuk. Méréseinkben tipikusan átlagosan 13 mérőpont vesz sikeresen részt.

5.3.2.1. Az algoritmus futásiidejének vizsgálata

Becslő algoritmusunkkal szemben gyors konvergenciát követelünk meg. Terünk a közeljövőre, hogy a folyamatosan frissülő mérési nyersadatok alapján a hálózat forgalmi térképe akár a mérés után kevés idő elteltével rendelkezésre álljon. A mérésiértékelés gyorsítására az egyik egyszerű észrevétel, hogy minden forrás

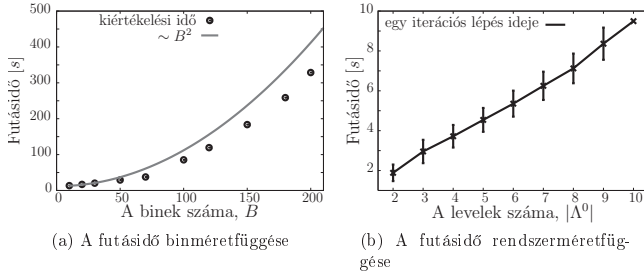
nyersadatai külön feldolgozhatók, így azokat több processzor párhuzamosan értékelheti ki. A gyorsítás érdekében bevezetett másik ötlet a következő. A nyersadatokat először nagy méretű binekkel kvantáljuk ($q \approx 100 \mu s$) és egyenletes eloszlást adunk meg *a priori* valószínűségeként, melynek következtében viszonylag gyorsan durva felbontású képet kaphatunk a belső szegmensek sorhossz-eloszlásairól. Következő lépésben az idősorokat egyre kisebb q kvantumokkal diszkrétizáljuk, és ennek megfelelően módosítjuk a Θ valószínűségek binelését is. A finomabb felbontáson elvégzett kiértékelések előtt az eloszlásokat véletlen zajjal perturbáljuk, hogy a likelihood-függvényt az esetleges lokális maximumából kibillentjük. Az 5.2. táblában megvizsgáljuk, milyen hatással van az algoritmus futásidejére a helyes kezdeti paraméterválasztás. Különböző binméret mellett összehasonlítottuk, mennyi időt vesz igénybe egy kiértékelés amikor egy durvább kiértékelésből származó paraméter beállítás lehetséges, szemben az egyenletes eloszlásról indított iterációkkal. Az 5.2. táblázatban a két futásidő hányadosát mutatjuk be. A helyes kezdeti paraméterválasztás akár fele akkora futásidőt jelent, melynek mértéke főleg a finom felbontású kiértékelésnél jelentős. Ez a tény rámutat arra is, hogy a kezdeti inicializálásra használt egyenletes eloszlás helyett célszerű áttérnünk arra, hogy minden szegmens kezdeti eloszlását az obszervatóriumban korábbi mérésből elérhető adatokkal töltsünk fel.

Binek száma	Gyorsulás
50	0,704
100	0,606
200	0,599
400	0,439

5.2. táblázat. Az előzetes tudásnak köszönhető futásidő csökkenésének mértéke.

A graduálisan finomodó felbontásokra elvégzett ismételt kiértékelésének módszere számos kellemes tulajdonsággal bír:

- A kiértékelés bármikor megszakítható, és szükség esetén tovább finomítható egy előző állapot Θ valószínűségeit választva kezdeti feltételként anélkül, hogy az algoritmus összes belső változóját mentenünk kellene.
- Kisebb felbontás mellett a kiértékelésben kevesebb mérési mintát elegendő felhasználni, hogy megbízható eredményt kapjunk.
- Minden kiértékelés után ellenőrizhető, hogy teljesül-e a konvergencia kritérium.
- Különböző szegmenst kiragadva lehetséges az eloszlások paramétereinek binméret függését megvizsgálni, amiből a számolt paraméterek hibájára adhatunk mértéket.



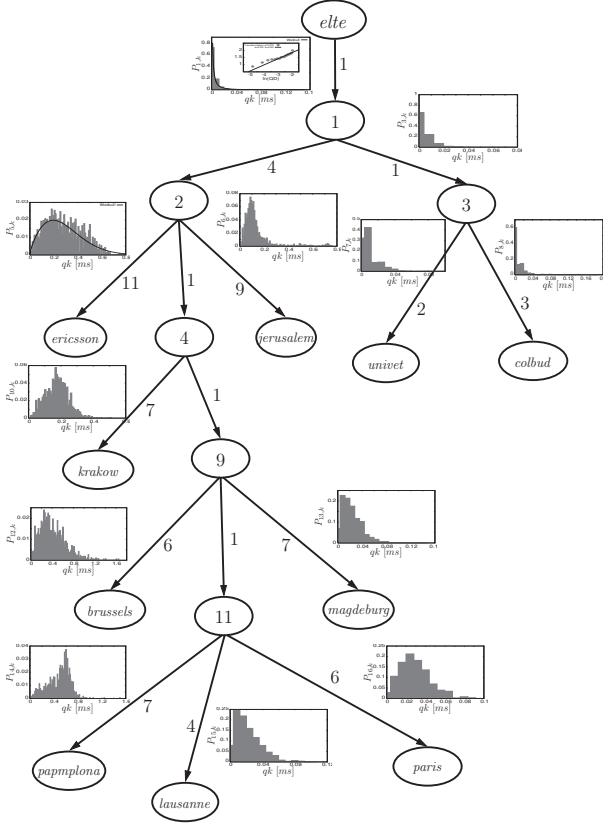
(a) A futásidő binméretfüggése

(b) A futásidő rendszerméretfüggése

5.13. ábra. A *message passing* módszerre alapozott becslés performancia vizsgálatát a *colbud.etomic.org* forrás egyik tomográfiás mérésére vizsgáltuk meg. (a) Az idő-felbontást, azaz a binek számát (B -t) növelve az algoritmus futásideje az $\mathcal{O}(B^2)$ alatt marad. (b) Megvizsgáltuk a vizsgált hálózat mérete, hogyan befolyásolja az algoritmus skálázását, és tapasztalatunk szerint a levelek számával egyenes arányban nő minden iterációs lépéshez szükséges idő.

A binméretek helyes megválasztásával a teljes kiértékeléshez tartozó futásidő jelentősen csökkenthető. A számításigény szempontjából legköltségesebb lépés az iterációkban felmerülő összegzések végrehajtása, melyek végső soron az 5.10. egyenlet implicit megadott konvolúcióját számolják. Ez $\mathcal{O}(B^2)$ skáláz, így a futásidő s iterációs lépését követően sB^2 nagyságrendű. Előre megválasztott q_i kvantálások B_i sorozata mellett legyen s_i iterációs lépés szükséges az i -ik kiértékeléshez. Exponenciális csökkenő binméretek választása mellett a teljes futásidő $s_1B_1^2 + s_2B_2^2 + \dots + s_kB_k^2$ lassabban nő, mint sB_k . Ezt támasztja alá az 5.13a. ábra.

Annak érdekében, hogy meggyőződjünk arról, hogy a fákra kidolgozott tomográfiás algoritmus képes lehet-e a jövőben több mérőpontot felölölő hálózati sorhossz tomográfiás becslését ellátni, numerikus vizsgálatokat folytattunk. Az 5.13b. ábrán megmutatjuk, hogyan skáláz a futásidő a rendszer méretével. A kiértékelés során a mérésben részt vevő nyelők közül módszeresen elhagytunk párat, azaz több forrás és a lehetséges 2–10 számú nyelők közötti összes lehetséges kombinációban elvégeztük a tomográfiás kiértékelést. Ezek közül a *colbud.etomic.org* forrás kiértékelését kiragadva az 5.13b. ábrán észrevevesszük, hogy a futásidő lineárisan nő a kiértékelésben részt vevő nyelők számával, ami a módszer jó skálázási tulajdonságait megerősíti.



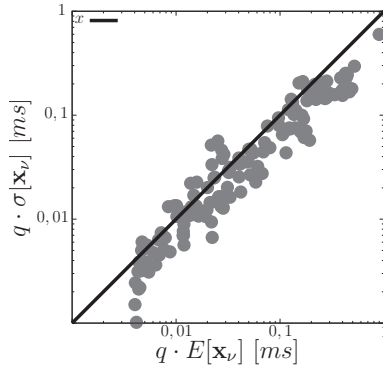
5.14. ábra. A 2007. január 11. 11:56 mérés egy tipikus forgalmdinamikát leíró statisztikákkal felruházott fája. Az éleken feltüntetünk a szegmens hosszát. Azokon a vonalakon, ahol nem tüntetünk fel eloszlásokat, $P_{v,0} = 1$ az eloszlás, azaz üres sorokat találunk, vagy olyan kis idők jellemzik a sor dinamikáját, amit nem tudunk felbontani. Ezek az elágazási pontok a gyors (≈ 10 Gbps) Akadémiai Hálózatban üzemelnek.

5.3.2.2. Analízis

A nagy precizitással jellemezhető mérések és az előző szakaszokban bemutatott gyors becslő algoritmus együttes használatával lehetővé tettük, hogy a

hálózat belsejét jellemző forgalmi állapot paramétereit nagy felbontással határozhatjuk meg.

Egy 11 mérőpontot tartalmazó tomoográfias mérés felbontott fáí közül az 5.14. ábrán példaként bemutatunk egyet, amiben az egyetemen működő (*elte.etomic.org*) végpont volt a forrás. A mérőpontok között kifizülő fa topológián az éleken feltüntetjük azok hosszát, ami definíciónk szerint a szegmensen belül felbont-hatatlan vonalakat méri. Amennyiben a szegmens hossza 1, akkor valódi vonalat tartunk fel, ha 2, akkor egy útvonalválasztó található az elágazási pontok között, stb. Az élek mellett feltüntetjük a rájuk jellemző sorban állási idők valószínűség-sűrűségeit is, melyek közül kettőt alaposabban megvizsgálunk. A sorhossz-eloszlások feltárt finomabb szerkezete lehetőséget teremt a sűrűségfüggvény alakjának további analizisére. Az ábrán az élek egy része mellett nem jelöljük a sorhosszakat, mert ezek a szegmensek a nagy sebességű *GÉANT* Akadémiai Hálózat gerincén vannak, és az elérhető időfelbontás mellett „üresnek” látszanak sorai, a triviális ($P_{\nu,0} = 1$) eloszlást számoljuk rájuk.



5.15. ábra. A megbecsült sorbanállási idők szórásai az átlagok függvényében. Az egyenes képlete $\sigma[\mathbf{x}_\nu] = E[\mathbf{x}_\nu]$. A szórásértékek a $0,5E[\mathbf{x}_\nu] < \sigma[\mathbf{x}_\nu] < 1,1E[\mathbf{x}_\nu]$ intervallumba esnek. Ettől nagyobb eltérést csak a nagy sebességű, azaz az alacsony várható értékű tartományban tapasztalunk, ami közel van a mérési felbontás nagyságához.

Az európai Internet hálózatot jól lefedő mérésünk alapján, a 13 forráshoz tartozó 183 sorhossz eloszlását együttesen vizsgálva az 5.8. ábrához hasonlóan most is három nagyságrendet felölölnek a sorban állási idők várható értékei. Az 5.15. ábrán a sorhossz eloszlások szórását ábrázoljuk az átlagok függvényében. Az átlag függvényében a szórás értékek széles tartományon lineárisan változnak, ami

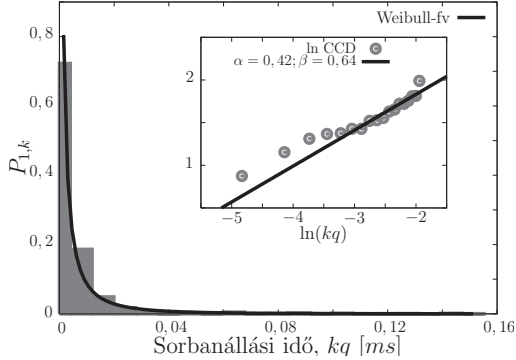
arra utal, hogy a sorhossz eloszlások egy függvényosztályba sorolhatók, „kicsit” eltérő alakkal. Az egyedi sorhossz eloszlások jól követik a Weibull-függvényeket, melyek várható értéke és szórása között teljesül a

$$\sigma[x_\nu] = \gamma(a)E[x_\nu],$$

összefüggés, ahol a a Weibull-függvény alak-paramétere, és az arányossági tényező

$$\gamma(a) = \frac{(\Gamma(1 + 2/a) - \Gamma^2(1 + 1/a))^{1/2}}{\Gamma(1 + 1/a)}.$$

A különböző sorhossz eloszlások illesztéséből az alakparaméter az $a \in [0, 5; 2, 0]$ intervallumba esik, amiből az arányossági tényező értékére $\gamma \in [0, 5; 1, 1]$ adódik, ami alátámasztja az 5.15. ábra adatait.



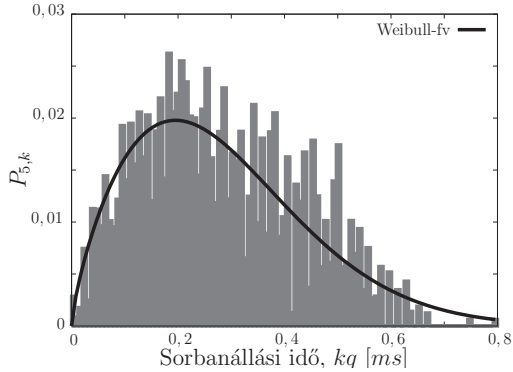
5.16. ábra. Az *elte.etomic.org* végpont vonalán elhelyezkedő első kimenő sor (az ún. *up-link* sor) várakozási idő eloszlását és az illesztett Weibull-függvényt ($\mathcal{P}(\mathbf{x}_1 > x) = \exp(-bx^a)$) mutatjuk be. A belső ábrán a kiegészítő eloszlás logaritmusa és a transzformált Weibull-függvény látható, ami hatványfüggvény. A dupla-logaritmikus skálán illesztett egyenes meredeksége kapcsolatba hozható a Hurst-kitevővel ($\alpha = 2(1 - \mathcal{H})$), ezen a példán $\mathcal{H} = 0,79$.

A mérésiértékelés során, találunk olyan hálózati szegmenseket, amelyek valódi vonalaknak felelnek meg, azaz ahol egyedi sorok dinamikája vizsgálható, szemben az összetett szegmensekével, ahol több sor együttes viselkedése, a konvolváltsor figyelhető meg. Az 5.16. ábrán az *elte.etomic.org* mérőpont kimenő vonalát mutatjuk, ami valódi vonal. Megvizsgálva a vonalhoz rendelt egy-sor eloszlásának lecsengését, a keresztforgalom önhasznós volta ellenőrizhető a Hurst-kitevőn

keresztül. Az illesztett Weibull-függvény paramétereiből Norros-képlete [71] alapján határoztuk meg a keresztforgalom idősorát jellemző Hurst-kitevőt:

$$\mathcal{P}(\mathbf{x}_\nu > x) \sim \exp\left(-\frac{(C-m)^{2\mathcal{H}}}{2k^2(\mathcal{H})c_v m^2}x^{2-2\mathcal{H}}\right), \quad (5.16)$$

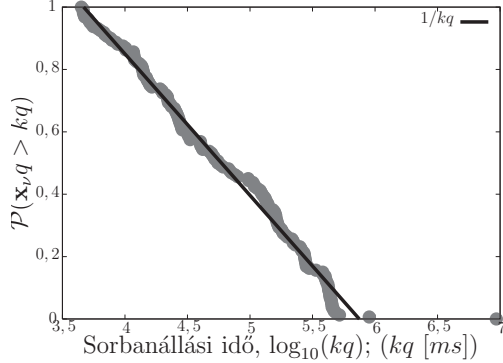
ahol C a vonal sebessége, m a keresztforgalmat jellemző átlagos ráta, $c_v = \sigma/m$ a keresztforgalom szórásának és várható értékének aránya, és $k(\mathcal{H}) = \mathcal{H}^{\mathcal{H}}(1 - \mathcal{H})^{1-\mathcal{H}}$, melyben \mathcal{H} a Hurst-paramétert jelenti. A hálózatkutatás irodalma szerint a forgalom önhasznó szerkezetére csak passzív mérési adatokon keresztül lehetett következtetni, melyek forrása egy-egy útvonalválasztó naplói voltak [72, 73].



5.17. ábra. A $\nu = 2$ és az *ericsson.etomic.org* csúcsok között található szegmens sorhossz eloszlása. A vonal tipikus példa a nagy sebességű maghálózat és a helyi 100 Mbps sebességű hálózat között található vonalakra. Az eloszlást Weibull-függvényel illesztettük.

Az 5.17. ábra egy összetett szegmens sorhossz eloszlására példa. Az egyedi sorhossz eloszlások konvolúciójaként előálló eloszlás nagyon hasonló formát követ, mint a végpontok között mérhető hálózati késleltetések eloszlásai. Ezt az eloszlást is jól illeszthetjük Weibull-függvényekkel, bár a keresztforgalmat jellemző Hurst-paraméter nehezen hozható összefüggésbe a Weibull-függvény paramétereivel.

Az elmúlt évtizedben az Internet skálázásával és fraktál jellegű tulajdonságával kapcsolatba hozható számos érdekes tanulmány született [17, 6]. Az Etomic mérési infrastruktúra nagy időfelbontása és a felgyorsított kiértékelő algoritmus azt is lehetővé tette számunkra, hogy az Internetet jellemző új skálázási törvényt fedezzünk fel. A sorhosszak várható értékének eloszlását vizsgálva érdekes skálázási tulajdonságot tártunk fel. Az 5.18. ábrán a 183 szegmensre jellemző várható



5.18. ábra. Az Internet egy újabb skálázási tulajdonsága. A 5.7. ábrához hasonlóan elkészített kiegészítő eloszlás fél-logaritmikus ábrázolása. Az illesztett folytonos vonal $\sim -\log(kq)$ eloszlást követ, ami a sorbanállási idők várható értékeinek $\sim 1/kq$ típusú skálázásának felel meg.

sorbanállási idők kiegészítő eloszlását ábrázoljuk fél-logaritmikus skálán, és az eloszlás farka hatvány-függvény szerint esik:

$$\mathcal{P}(E[\mathbf{x}_\nu] > x) \sim C_1 - C_2 \log(x),$$

ami arra utal, hogy a hozzá tartozó sűrűségfüggvény $\sim 1/x$ hatvány-függvény szerint skáláz. Ebbe az osztályba tartozó „sűrűségfüggvények” rendkívüliek, mert véges levágások hiányában nem létezik sem átlaguk, sem szórásuk, sőt nem is igazi sűrűségfüggvények, mert nem normálhatóak.

Jelenleg nem ismerjük a pontos okát annak, hogy milyen körülmények között és miért és áll elő ez a skálázási tulajdonság. Egy lehetséges magyarázat szerint a hálózatba épített kiszolgálási kapacitások elosztásával hozható összefüggésbe. A vonalakra jellemző sorbanállási idők közelítőleg

$$x \approx \frac{P}{C - m},$$

alakban írhatóak, ahol $C - m$ a szabad kapacitás mértéke. Azokon a vonalakon, ahol elhanyagolható a keresztforgalom m nagysága, a képlet tovább közelíthető P/C hányadossal. Gyanús, hogy a sorbanállási idők várható értékeinek $1/x$ típusú skálázásának hátterében valójában a szabad linkek $1/C$ skálázása mutatkozik meg.

5.4. A traceroute-tomográfia ötlete

Fás tomográfiás méréseink segítségével sikerült az Interneten működő egy-sor viselkedések rejtelmeibe is belepillantanunk. Az egyedi vonaltulajdonságok feltárása a hálózatkutatásban nagyon fontos feladat, olyasmi mint egyre jobb felbontású mikroszkópot építeni, amikor az anyagszerkezetet vizsgáljuk. Ideális esetben egy teljes útvonal mentén fel lehet tární a sorokat jellemző késleltetéseket. Az ötlet az útvonalfeltáró **traceroute** program és a csomagvonatküldő mérőprogram házasításával oldható meg.

A mérőprogramot megvalósítottuk, mely működésének első fázisában egy cél irányába feltérképezi az útvonalválasztók sorozatát. Ennek ismeretében a program a második fázisban olyan csomagvonatot küld a hálózatba, melynek címzettjei az érintett célpont, de a csomag életét szabályzó *TTL* mezői eltérnek. Az első csomag életkora az első útvonalválasztón jár le, a következő eggyel később, és így tovább. A mérés során a minták körüljárási idejeit mérjük, melyek gyakorlatilag egy fát járnak be. Az eredeti minták által bejárt útvonal ismert, valódi vonalakat jelent, az elhaló csomagokra érkező válaszüzenetek összetett szegmenseken érnek vissza a mérőponthoz, mely szegmensek ismeretlenek. A mérésben kifeszülő fa úgy néz ki, hogy a mérőgép alkotja a fa gyökerét és leveleit egyaránt.

Előzetes traceroute-tomográfiás méréseink még nem hozhatták meg az elvárt sikert. Az egy-sor vonalakra eredményül kapott eloszlások tartói nagyságrendekkel felülmúlták várakozásainkat, az átlag *ms* nagyságrendjébe esik szemben az elvárt $10\ \mu\text{s}$ -mal. Az anomália mögött meghúzódó prózai magyarázat abban rejlik, hogy a feltárt sorbanállási időbe beleszámolódik az útvonalválasztó válasz-csomag felépítésével eltöltött ideje, amelyet ma még nem ismerünk kellően, hogy kikomponáljuk az összetett becslési eredményekből.

5.5. Diszkusszió

A hálózat dinamikus forgalmának pontos ismerete nemcsak a modellalkotás szempontjából fontos tudás, hanem erre alapozva a forgalomvezérlés eddig ismeretlen új elvei alapozhatók meg. A hálózat állapotváltozói közül fontos kiemelni a sorok kiszolgálási idejét, melyet gyakorlatban a valódi hálózatban közvetlenül nem mérhetünk. A sorbanállási idő várható értéke és ingadozásai számot adnak a forgalom helyi torlódási viszonyairól.

Ebben a fejezetben aktív mérésen alapuló tomográfiás módszereket mutatunk be, melyek segítségével a sorbanállási idő becsülhető. Tomográfiás méréseinket az Etoomic mérőhálózat segítségével az európai Interneten végezzük, a feltárt belső paraméterek számot adnak a hálózat pillanatnyi forgalmi állapotáról. Kísérleteink alapján megállapítottuk, hogy a sorbanállási idő valószínűségi változói Weibull-eloszlásokat követnek, aminek alakparamétere függ a szegmens méretétől, egy-sorok esetén a keresztforgalom fraktál tulajdonságáról is számot ad. A

nagyságukban 3 rendet átölelő várható értékek együttes eloszlása bizonyos forgalmi körülmények között skálázási tulajdonságot mutat.

Érdemes megjegyeznünk, hogy a fejezet sorhossz tomográfiás módszerei szignifikáns csomagvesztéssel bíró hálózatban átalakíthatók csomagvesztést becslő modellre, csak az 5.1. képletben szereplő változókat kell átértelmeznünk. Erre az esetre Y a minta célba érkezési valószínűségnek logaritmus, és X a minta áthaladási valószínűségének logaritmus.

6. fejezet

Torlódási hullám kialakulása egy absztrakt hálózatmodellben

6.1. Motiváció

A különböző kommunikációs hálózatok forgalmának modellezése fontos kutatási ág. Számos érdekes jelenséggel találkozunk a méréseket és a szimulációs eredményeket vizsgálva. Antoniou szerzőtársaival [74] hasonlóságot tár fel a hálózati csomagforgalom és a granuláris anyagok csövekben történő áramlása között. A kihasználatlan kapacitásokért folytatott küzdelemben a csomagforgalom szabályzására jellemző gyorsítási és lassítás folyamatok nagyon hasonlítanak az autós közlekedésben is tapasztalható jelenségekhez. Ezeket sokszor sejtautomata modellekkel próbálják megragadni [75, 76, 77, 78], és számos analógiát tártak már fel az irodalomban, statisztikai jellemzőket megfelelően egymásnak [79, 80]. Mindkét rendszerre jellemző az önhasonló viselkedés [16, 17, 34] valamint a torlódási rezsimben megfigyelhető az $1/f$ -zaj jelenléte [81, 19, 20]. Akár a granuláris anyagok áramlását, akár az autós forgalommodelleket tekintve, az egyik legérdekesebb jelenség a sűrűsödési hullámok kialakulása és terjedése [82, 83, 84]. Az autós forgalommodellekben a fékezés és gyorsulás váltakozásának következtében kialakuló dugóhullám a menetiránnyal szemben halad. Fukuda és szerzőtársai [85] az Interneten működő útvonalválasztó egységek a valódi áthaladó forgalomról készített eseménynaplót elemezve, a torlódások elmozdulását mérték ki több *router* között. A torlódás mértékéről impliciten számot adó, körülfordulási idők térbeli és időbeli korrelációit megvizsgálva arra a következtetésre jutottak, hogy az élő hálózat terhelte útvonalválasztóiról a szomszédos üres útvonalválasztókra tevődik át a torlódás.

A kialakuló torlódáselmozdulási jelenség mögött a forgalomforrások csatolt dinamikája tehető felelőssé. Célul tűztük ki, hogy belássuk, egy csupasz, – a hálózat bonyolultságától eltekintő, – modell is képes érdemben számot adni erről a jelenségről. Ebben a fejezetben egy könnyen tárgyalható modellrendszerben kialakuló

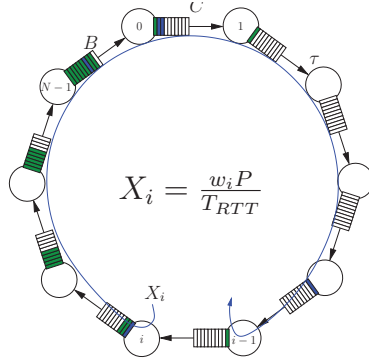
forgalommintázat tanulmányozásával foglalkozunk. A modellben azonos típusú útvonalválasztó egységekből és azonos irányban irányított vonalakból felépített hálózatot valósítunk meg az ns hálózatszimulátor segítségével [25]. Hálózatválasztásunk egyik oka a könnyebb tárgyalhatóság, valamint az, hogy ehhez hasonló hálózatfelépítéssel dolgoznak számos, a granuláris anyagok áramlását leíró modellben, s így lehetőségünk nyílik a modellek analógiáit feltárni és az eredményeket összehasonlítani. A periodikus határfeltételt megvalósító gyűrű-topológia lehetőséget teremt arra, hogy a kialakuló dugóhullámokat jól elszigetelten, tisztán tanulmányozhassuk anélkül, hogy figyelembe kellene vennünk a véges méretből adódó szélhatásokat, inhomogenitásokat, illetve a valódi Interneten megvalósuló komplex topológiát.

A 6.3. fejezetben bemutatjuk, hogy a vizsgált rendszer magától kritikus dugulási állapotba fejlődik. A kialakuló stacionárius állapotot jól jellemzi, hogy mind a dugult útvonalválasztó egységek gyűrűben elfoglalt helye – azaz ahol a pillanatnyilag tapasztalható csomagvesztési mutató magas – mind a csomagtovábbítási aktivitás – azaz az időegység alatt vonalra helyezett csomagok száma – időben mozog, a tényleges csomagforgalom irányával szemben halad, dugulási hullám alakul ki. A kialakuló dugóhullámot leíró profil egyenletét előállíthatjuk a csomagtovábbítási aktivitások ismeretében, és mérhetővé válik a szimulációs elrendezésben kialakuló hullám sebessége. A 6.4.1. fejezet végén absztrakt modellt alkotunk, ami a benne levezetett képleteken keresztül számot ad a szimulátorban tapasztalható dugóhullám alakjáról és terjedési sebességéről.

6.2. A gyűrű alakú hálózat

Az általunk vizsgált modellben kezdetben különböző N darab egymással minden tulajdonságában megegyező útvonalválasztó-csomagtovábbító egységet kapcsolunk össze úgy, hogy azok gyűrű alakú hálózatot formáljanak. Az információ-csomagok az óramutató járásával megegyező irányban haladnak, ahogy az a 6.1. ábrán látható. A szimulációs környezetben minden útvonalválasztó egységhez egy a csomagforrás és a csomagnyelő szerepét betöltő *terminált* kötünk, ezek generálják a hálózati forgalmat. Az útvonalválasztókat összekötő vonalak kapacitása ugyanakkora C , – és állandó a rájuk jellemző τ propagációs idő is. Az útvonalválasztó egységekbe befolyó csomagok keveréket alkotnak. Egy részük a közvetlen mellé kapcsolt forrás *terminálról* érkező csomagokból áll, míg a többi csomag a háttérforgalom része, ami valójában a többi forrás forgalmából tevődik össze. A befolyó csomagok árama időben változó függvény és esetenként meghaladja a kimenő vonal C kapacitását. Annak érdekében, hogy ilyen esetekben megelőzzük a csomagok elveszését, minden útvonalválasztóhoz B csomag befogadására és ideiglenes tárolására alkalmas puffert rendelünk.

A forrás egységek egyszerű szabály szerint küldik csomagjaikat: mindegyik az óramutató járásával ellentétes irányban számolt legközelebbi szomszédnak küld



6.1. ábra. A szimulátorban megvalósított gyűrű alakú hálózat. Minden lehetséges i , $(\forall i : 0 \leq i < N)$ helyen egy folyamatosan forgalmazó alkalmazás üzemel. Az útvonalválasztó egységek legnagyobb kimenő kapacitása C . A vonalakat állandó τ propagációs idővel jellemezzük. A csomagok tárolására véges B méretű pufferek vannak az útvonalválasztókba építve. A forgalom az óramutató járásával megegyező irányban halad, az i pontban eredő folyam az $i - 1$ helyen nyelődik el. A szimuláció során beállított paraméterek többsége az Internetre is jellemző értéket veszi fel. ($C = 10^7$ bits/s, $\tau = 0,031$ s, $B = 300$ csomag, $P = 4416$ bit, $N = 10$.)

folyamatosan, így az irányított vonalak segítségével a leghosszabb bejárható utat feszítjük ki az adott topológián. A megvalósított forgalom granuláris, hiszen azt mikroszkopikusan véges P méretű diszkrét P méretű csomagok valósítják meg. Az egyedi csomagküldés dinamikáját a 2.3. TCP/IP protokoll írja le. A protokoll szerint megvalósított csomagküldési sebesség szabályzás additív növekedéssel és vesztés esetén a multiplikatív csökkenéssel adaptálja a forgalom nagyságát a rendszer kínálta szabad erőforrásokhoz.

A dinamikát a 2.1. és a 2.2. mozgásegyenletei helyett a csomagküldési sebességre vonatkozó egyenletekkel írjuk le. A csomagküldés sebesség X , az egy körüljárási idő alatt kiküldött csomagok számát adja meg, $X = Pw/T_{RTT}$ hányadossal közelítve. Modellünkben a körüljárási idő skáláján eltekintünk annak változásától, így folytonos közelítésben a csomagküldési sebességet jól leírja az

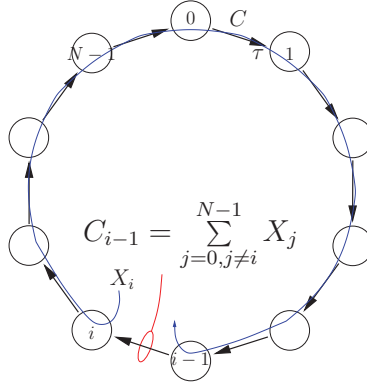
$$\frac{dX}{dt} = \frac{P}{T_{RTT}^2(t)}, \quad (6.1)$$

$$X(t_+) = \beta X(t_-) \quad \text{/csomagvesztés esetén/} \quad (6.2)$$

egyenletpár, amiben, hűen követve a TCP tipikus megvalósítását, $\beta = 1/2$ para-

méterválasztással élünk.

6.3. Szimulációs eredmények



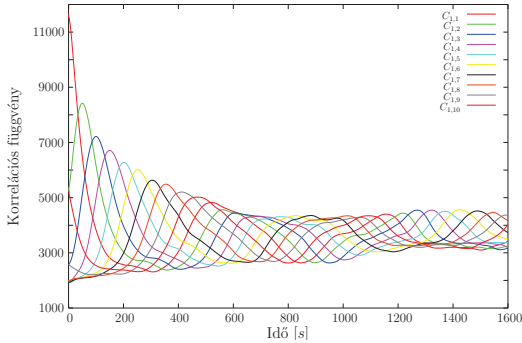
6.2. ábra. A folytonos közelítés modellje. Az i -ik forrást jellemző küldési sebesség X_i . Az $i-1$ és i helyet összekötő vonalon áthaladó folyamokhoz tartozó küldési sebességek összege adja meg a pillanatnyi kiszolgálási sebességet.

Ebben a szakaszban bemutatjuk az ns hálózati szimulátor segítségével végzett kísérleteink eredményeit. A hálózati összeállítás geometriáját és paramétereit a 6.2. ábrán vázoljuk. A lényeges paraméterek megválasztása során törekedtünk arra, hogy azok megegyezzenek, vagy értékük közel essen a valódi hálózatokon tapasztalható értékekkel. Először a legegyszerűbb esetet vizsgáltuk meg, amikor minden végpont végtelen ideig küldi hálózatba csomagjait, ami gyakorlatban a végtelen hosszú adatfájl letöltésének felel meg.

A 6.3. diagram a szimulációs hálózatban megjelenő torlódási hullám megjelenését és terjedését ábrázolja, $N = 10$ választott rendszer méret mellett. A diagram vízszintes kiterjedése az idő múlását követi, a függőleges irány a terminálokat számlolja, a színezés árnyalata a küldési sebesség nagyságát érzékelteti. Jól látható, hogy a szimulációjának kezdetekor rövid, körülbelül 500 s ideig tartó, transziens viselkedés tapasztalható, majd stabil mintázat alakul ki, ami az óramutató járásával ellentétes irányban mozog. A transziens hossza a rendszer méretétől és a véletlenül kiosztott kezdeti feltételektől függ. Ezután figyelmünket a kialakult torlódási mintázat vizsgálatára fordítottuk.



6.3. ábra. A dugóhullám térbeli és időbeli viselkedése az $N = 10$ rendszerméret mellett. A vízszintes tengely 2000 s-nyi időt ölel fel, a függőleges tengely az i helyet határozza meg. Meg kell jegyezni, hogy az alsó sor ($i = 0$) és a felső sor ($i = 9$) szomszédosak. Az ábra árnyalata a küldési sebesség (X_i) nagyságát érzékelteti, a világos foltok a torlódás következtében beállt kis sebességet jelentik, a sötét foltok $\approx 700P/T_{RTT}$ sebességnek felelnek meg. A tranzienst követően jól látszik, hogy a forgalommal szemben halad a dugóhullám.



6.4. ábra. Az $N = 10$ méretű rendszerben kialakult stationer állapotban képezzük X_1 és X_i ($i \in [1, N]$) küldési sebességek korrelációs függvényét, amely csak akkor alkalmas a hullám sebességének meghatározására, ha a hullámforma tiszta és stabil. A korrelációs függvények lecsengő jellege utal arra, hogy még stationer állapotban sem tökéletesen egyforma minden periódusban a hullámalak.

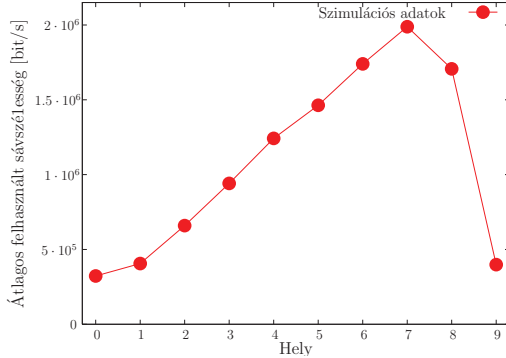
Hosszabb szimulációs futtatásra, a kezdeti bizonytalan adatoktól megtisztított $X_i(t)$ idősorok között meghatároztuk a $C_{1,i}(t) := \int X_1(T-t)X_i(T)dT$ korrelációs függvényt, amit a 6.4. ábrán mutatunk be. A hullámfront sebességét leolvashatjuk a korrelációs függvény csúcshelyei alapján. A módszer hátránya, hogy csak tiszta, szemmel is szépen felismerhető hullámfrontok esetén alkalmazható, számításigényes és pontatlan. A kiragadott példában a korrelációs függvény mód-

szerével $\approx 0,02 \text{ s}^{-1}$ sebesség adódik, a hullámok kézzel elvégzett leszámolásával $\approx 0,016 \text{ s}^{-1}$ sebességet határozhatunk meg. E módszereknél elegánsabb, kifinomultabb megoldást kerestünk.

A stacioner állapotban kialakuló torlódás nagyban hasonlít az autóforgalomban tapasztalhatóra. A torlódási hullám sebessége nagyjából állandó, átlagát könnyen meghatározhatjuk a hullámforma súlypontjának átlag sebességének meghatározásával. A súlypont helyének és a belőle származtatott sebesség meghatározásánál körültekintően kell eljárunk, a periodikus határfeltételt figyelembe kell vennünk. Legegyszerűbb a komplex egységkör N -ik egységgyökeire vetíteni az egyes terminálok pillanatnyi csomagküldési sebességét, abból komplex számsorozatot képezve. A számsorozat komplex átlaga közvetlenül megadja a hullámforma súlypontjának pillanatnyi helyét. Tehát a súlypont helyéről a

$$\langle i \rangle(t) = \frac{N}{2\pi} \arg \left(\sum_{j=0}^{N-1} X_j(t) e^{i(2\pi/N)j} \right)$$

kifejezés ad számot. A dugóhullám terjedési sebessége ennek a kifejezésnek időderiváltja.



6.5. ábra. A mozgó hullám alakja, melyet az együttmozgó koordináta-rendszerben elvégzett időátlagolás után látunk.

A hullámprofil sebességének meghatározása után lehetőségünk nyílik meghatározni annak alakját. A dugóhullám pontos alakja időben változik, különböző szimulációs időpontban megvizsgálva más-más formát ölt, ezért a küldési sebesség hullámfrontjára átlagot képzünk. Az együttmozgó koordináta-rendszerben a küldési sebesség képlete: $X_{i' + \lfloor \langle i \rangle \rfloor}(t)$, ahol i' a helyet a hullámforma súlypontjától méri, ezzel mintegy befagyaszttjuk a hullám propagációját. A hullámprofil a

„befagyasztott” hullámformák időátlagával határozzuk meg, melynek eredménye a 6.5. ábrán szerepel.

6.4. A hullámfront kialakulása

Ugyan az eredeti TCP-dinamika mozgásegyenleteinek folytonos modellel közelítése durva becslés, a kialakuló hullámforma legfontosabb tulajdonságairól pontosan számot ad. Ehhez a csomagvesztés folyamatára vonatkozó következő feltetéssel kell élnünk.

Az $i - 1$ és az i helyen elhelyezett útvonalválasztókat összekötő vonalon a kihasznált vonalkapacitás $C_{i-1}(t)$. Ez a mennyiség ezen a vonalon jelen levő egyedi információfolyamokhoz tartozó terminálok csomagküldési sebességeinek összegeként áll elő. Modellünkben ez pont az összes végpont forgalmának együttese, leszámítva az i -től $i - 1$ -ig küldő terminál csomagforgalmát, azaz:

$$C_{i-1}(t) = \sum_{j=0, j \neq i}^{N-1} X_j(t) = \sum_{j=0}^{N-1} X_j(t) - X_i(t), \quad (6.3)$$

ahol a periodikus határfeltételt a $i = N \Leftrightarrow i = 0$ azonos megfeleltetéssel teljesítünk. A nyugtacsomagokból származó forgalom nagysága – ami az $i - 1$ -ik végponton ered és rögtön i -ben elnyelődik – elhanyagolható, a kis csomagméret és a teljes forgalomhoz viszonyított kis hozzájárulás miatt. Ezt illusztrálja a 6.2. ábra.

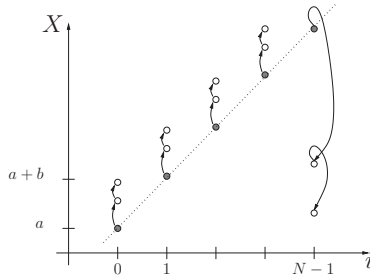
A modellrendszerben akkor alakul ki a csomagvesztéssel járó torlódás, amikor a kihasznált kapacitás $C_i(t)$ meghaladja a vonal C sávszélességét. A modell 6.3. egyenletei szerint a legnagyobb vonalkihasználtság az $i = i^* - 1$ helyen tapasztalható, ahol i^* -re jellemző, hogy itt a legkisebb a csomagküldési sebesség $X_{i^*}(t)$. Azt kell megvizsgálni, hogy az $i^* - 1$ -ik útvonalválasztóból eredő vonalon, melyik folyamból veszik csomag. Ezen a vonalon elvben bármelyik folyam csomagja elveszhet, vagyis egyedül a i^* helyen található terminál forgalma maradhat érintetlen. A szimulációs kísérleteink gyakorlati tanulsága szerint majdnem kivétel nélkül abból a folyamból veszik csomag, melynek forrása a vonal elején van, vagyis amelyeknek csomagküldési sebessége X_{i^*-1} . Ennek oka abban rejlik, hogy a várakozási sorokat kiszolgáló egység a kisméretű nyugtákat sokkal gyorsabban dolgozza fel, mint a valódi információt hordozó nagyméretű csomagokat, amik visszaérve a feladóhoz, egyszerre hosszabb csomagsorozatokat, *burstöket* indukálnak. Mintegy pozitív visszacsatolást megvalósítva, a nyelőkhöz burstökben megérkező információk csomagokra a nyugták is sűrűn keletkeznek, és hamar visszajutnak a forráshoz, így a forrás végpontok még inkább burstösen küldik újabb csomagjaikat. Az öngerjesztő folyamat eredményeként az a forrás, amelyik közvetlenül egy közel telt pufferű sorba küldi burstjeit, sokkal nagyobb valószínűséggel veszít.

Az imént vázolt mechanizmus a felelős a rendszerben megfigyelhető torlódási mintázat kialakulásáért. Az $i^* - 1$ helyen található terminál folyamatosan csomagvesztéseket szenved mindaddig, míg saját küldési sebességét meg nem haladja a következő leggyorsabb, azaz mindaddig veszít, amíg nem teljesül a $X_{i^*-1} > X_{i^*}$ feltétel. Ettől a ponttól számítva X_{i^*-1} lesz a legkisebb sebesség a rendszerben, a C_{i^*-2} vonalkihasználatosság lesz a legnagyobb, és a $i^* - 2$ helyen működő forrás veszi át a veszítő szerepet. A gondolatmenet szerint a dugulás egy hellyel elmozdult óramutató járásával ellenkező irányba. A hullám néhány teljes kör megtétele után már a 6.5. ábrán vázolt hullámfrontot mutatja.

6.4.1. A hullámfront burkolója

A hullámfront alakja fűrészfogszerű, egyenes felfutású a végeket meredek eséssel összekötő burkolóval írható le. Ebben az alfejezetben a hullámfront burkolóját jellemző egyenes paramétereit vezetjük le modellünk segítségével. Kezdetben tegyük fel, hogy az egyenest alkotó pontok az $i \in [0; N - 1]$ tartomány egész értékein vannak értelmezve, és legyen a legkisebb küldési sebesség a 0-ik helyen tapasztalható:

$$X_i(0) = a + bi. \quad (6.4)$$



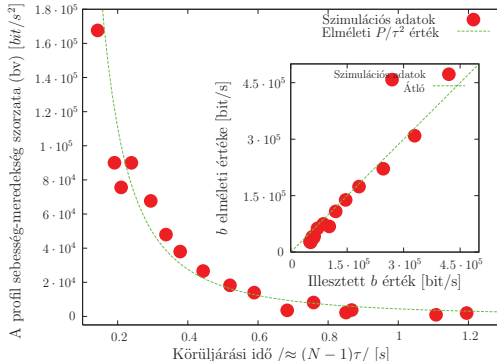
6.6. ábra. A küldési sebesség időfejlődése. $N = 5$, $l = 2$, X_4 mindaddig csökkenti értékét, amíg az az X_0 értéke alá nem esik.

A 6.6. ábrán szemléltetésül egy kisebb, $N = 5$ méretű rendszerre mutatjuk meg a hullámprofil időfejlődésének lépéseit. Az előző szakaszban vázolt gondolatmenet alapján, a kezdeti feltételek miatt a $N - 1$ helyen következik be az első csomagvesztés. Óránkat ehhez a pillanathoz állítjuk, és attól az időponttól kezdjük a tárgyalást, amikor a $C_{N-1}(0) = \sum_{i=0}^{N-1} X_i(0) - X_0(0) = C$ kifejezés teljesül.

Ezen kezdeti feltétel a hullámprofil kezdeti alakjára kirója, hogy

$$C = (N - 1) \left(a + \frac{N}{2} b \right). \quad (6.5)$$

A csomagvesztés után közvetlenül a $N - 1$ terminál csomagküldési sebessége multiplikatív csökken $X_{N-1}(0+) = \beta X_{N-1}(0-)$, míg az összes többi csomagküldési sebesség érintetlen marad. Ezt követően a következő csomagvesztésig mindegyik csomagküldési sebesség uniform módon egyenletesen növekedik $X'_i = X_i + (1-\beta)X_{N-1}/(N-1)$ mértékben. Az első és a második csomagvesztés között a $N - 1$ helyen a sebesség $X'_{N-1} = qX_{N-1}(0)$ szerint változik, ahol a tényező értéke $q = \beta + \frac{1-\beta}{N-1}$. A folyamat l -szer ismétlődik meg, mígnem a $q^l X_{N-1} \leq X_0$ feltétel teljesül. Az l veszteségi folyamat eredményeként a profil egy hellyel balra mozdul, megtartva annak lineáris alakját, és b -vel növelve, X_{N-1} kivételével, az összes jellemző küldési sebességet. A 6.1. egyenlet segítségével kiszámolható, hogy a hullámfront egy hellyel balra lépéséhez vezető folyamat $T_p = bT_{RTT}^2/P$ időt igényel, amiből ered a hullámfrontot jellemző $v := 1/T_p$ sebesség, melynek természetes mértékegysége $1/s$.



6.7. ábra. Az elméleti képletek tesztelése. A szimulációs környezetben különböző rendszerméret viselkedését vizsgáltuk meg ($N = 5 \dots 20$) és ellenőriztük, hogyan teljesül a $v b T_{RTT}^2 / P = 1$ összefüggés. A kiértékelések során adatainkból kivágtuk a tranzienst szakaszokat, majd 5000 s hosszú mintákra meghatároztuk b , T_{RTT} és v értékeit. (részábra): az (6.6) kifejezés ellenőrzése. A b paraméter átlagait ábrázoljuk az elméleti képletünk alapján (Mj. q^l értékét a szimulátor szolgáltatja.)

A levezetett formulából megbecsült sebességértékeket összevetettük az ns hálozáti szimulátor mérési adataira kiszámolt sebességekkel. A szimulátorban gyűjtött adatokból előállítottuk az átlagos körüljárási időt, a hullámprofil alakleíró

paramétereit valamint annak sebességét. A 6.7. ábrán jól látszik, hogy a mérési eredmények jól követik a $vbT_{RTT}^2/P = 1$ azonosságot.

A 6.5. és a 6.4. képletek alapján kifejezhető a hullámprofil meredeksége:

$$b = \frac{2C(1 - q^l)}{(N - 1)(N - 2)(q^l + 1)}, \quad (6.6)$$

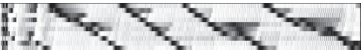










ahol l meghatározása külön feladat. A felvázolt matematika modell számos pozitív l értéket megenged, aminek felső korlátját b értéke határozza meg. Numerikus szimulációs tapasztalataink alapján, a legnagyobb megvalósítható l érték stabil a hullámfront kis perturbációival szemben. Előállítottunk olyan kezdeti feltételeket, amik folyományként kisebb l értékek valósultak meg, ám a hullámfront stacioner állapotában bármilyen perturbáció hatására a rendszer egy nagyobb l -hez tartozó állapotban stabilizálódott. Az ns hálózati szimulátorban minden esetben a modell által megengedett legnagyobb l értékhez tartozó állapot valósult meg. A 6.7. ábrába ékelte grafikonon összehasonlítottuk a becsült profilmeredekségeket a hálózatszimulátoros mérések nyersadataira számolt meredekséggel. A kétféleképp meghatározott meredekség adatai jó egyezést mutatnak, modellünk konzisztens.

6.5. A dugóhullám stabilitása

Az eddig bemutatott szimulációs eredmények azt az egyszerű esetet tárgyalták, amikor minden terminál ideálisan végtelen hosszú adatfájlt tölt át. A valóság ennél sokkal színesebb, az igazi hálózatban hosszabb-rövidebb fájlokat töltenek le a felhasználók, egy-egy kiszolgáló végpontra párhuzamosan több kiszolgálási kérelem érkezik, a felhasználói magatartás újabb szabadsági fokot visz a komplex rendszerbe.

Egyszerű modellünkben megvizsgáltuk, milyen hatással van a kialakuló hullámfrontra a paramétereknek egy részének bevezetése, változtatása. Szimulációs kísérletünkben megengedtük, hogy egy végpontra párhuzamosan több (n) információs folyam éljen egymás mellett. Minden egyes folyam egy-egy *on-off*-folyamatként volt megvalósítva úgy, hogy a hallgatási szüneteket megadó valószínűségi változók azonos paraméterű exponenciális-eloszlásból lettek kisorsolva. Ezzel homogén Poisson-érkezési folyamatot írunk elő, ami az emberi kommunikáció kezdeményezésre, például a telefonálások között eltelt időkre kimutatható. A bekapcsolás időtartamát az átvitt adatmennyiség szabta meg, ami Pareto-eloszlást követett, amely eloszlás jellemző a személyi számítógépeken és a webkiszolgálókon elérhető fájlok méretére [86]. Ezzel a paraméterválasztással utánozzuk az Internet *world-wide web* forgalmát.

Az így felépített modellben is kialakul a dugóhullám, de a paraméterek megfelelő elhangolásával a haladó torlódásfront gyakorlatilag elnyomható. A különböző paraméterválasztás mellett, a rendszer bemelegítése után kialakuló állapotot a 6.1. ábráin szemlétetjük.

n	$\langle F_s \rangle$	$\langle t_{OFF} \rangle$	Időtartam 200 s
10	10^8	0,5	
10	10^7	0,5	
10	10^6	0,5	
10	10^5	0,5	
10	10^8	1	
10	10^8	5	
10	10^8	10	
10	10^6	1	
25	10^6	1	
50	10^6	1	
100	10^6	1	

6.1. táblázat. A modellben kialakuló forgalmi mintázat a *www*-forgalom különböző paramétereit mellett.

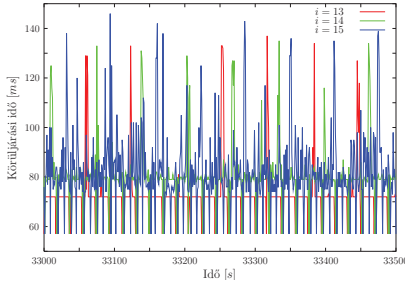
6.6. Diszkusszió

A 6.4. fejezetben megmutattuk, hogy számítógéphálózatokban az egyedi forgalomgeneráló dinamikák kölcsönhatásának folyamánként természetes módon kialakulnak torlódási mintázatok. A vizsgálat tárgyát képező egyszerű modellrendszerben a hullám kialakulásáért felelős mechanizmus oka a forgalomszabályzó dinamika tulajdonságaiban és a csomagvesztési mechanizmusban keresendő. A modell szerint csomagvesztés legnagyobb valószínűséggel a nagy sorokat tartó vonalakra közvetlenül kapcsolódó terminálok forgalmában tapasztalható. A dugult vonalon osztozó terminálok csomagküldési sebességüket tovább növelik, ezzel lehetőség nyílik arra, hogy a dugulás a szomszédos vonalra tevődjön át a megengedett csomagforgalommal szemben. Modellünkben a kialakuló hullám geometriájáról és annak sebességéről pontosan számot adó képleteket vezettünk le.

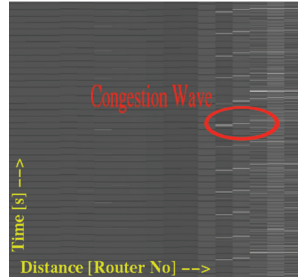
Megfontolásaink általánosak, előrevetítik, hogy bonyolultabb geometriában is

megjelenjen a forgalommal szemben haladó dugóhullám.

Elvárásunkat, hogy a valódi Interneten is tapasztalunk a torlódás elmozdulására utaló jeleket célzott kísérletekkel kívántuk igazolni. A 3.3.2. fejezetben részletesen bemutatott **traceroute** és **ping** alkalmazások segítségével egy Dél-Afrikai webszolgáltató felé térképeztük fel az utat jellemző körüljárási időket. Előzetesen feltártuk az egyetemről a szolgáltató felé vezető utat, majd az úton fellelhető összes útvonalválasztó irányában mértük folyamatosan a körüljárási időket. A mintegy fél napot felölelő mérésben begyűjtött idősoraiban felleltünk olyan szakaszokat, melyek a torlódás elmozdulására utalnak.



(a) Körüljárási idősorok



(b) RTT hely és időfüggése

6.8. ábra. (a) Az RTT mérésben részt vevő útvonalválasztók közül három szomszédos pontra jellemző idősor. Az ábrázolt idősorokon megfigyelhető ingadozások csatoltak, egymásutániságuk nem sérül. (b) Az RTT mérési adatok a teljes útvonalra és a mérés egy kiragadott időtartamára szürke skálán ábrázolva, a világos foltok felelnek meg a kiugrásoknak. A bekarikázott rész olyan mintát követ, amelynek az igazi hálózatban előforduló torlódási hullámot képzeljük.

A 6.8a. ábrán három darab sorban egymás utáni útvonalválasztó felé mért körüljárási idősort mutatunk be. Jól megfigyelhető, hogy esetenként kiugrások jelennek meg az idősorokban, és tipikusan a távolabbi routerhez tartozó idősor csúcsa tendenciózusan késik a többihez képest. A 6.8b. ábrán az körüljárási időtartamok hely és időfüggését mutatjuk meg, melyről leolvasható, hogy a forgalomintázatban, a távolabbi útvonalválasztókhoz tartozó idősorok ingadozásai csatoltak, ami jele lehet a torlódás elmozdulásának.

7. fejezet

Összefoglalás

A dolgozatban a számítógépes kommunikációs hálózatok forgalmi tulajdonságait vizsgáltuk kísérletező és elméleti fizikusi hozzáállással. Megmutattuk, hogy az Interneten elterjedt csomag alapú forgalom szabályzó dinamikával leírt ágensek a hálózat erőforrásaiért folytatott versengésükkel olyan csatolást visznek a rendszerbe, melynek hatására a hálózatban érdekes torlódási mintázatok alakulnak ki.

Célul tűztük ki, hogy mérési eljárásokat dolgozzunk ki és honosítsunk meg, melyek segítségével az Internet hálózatban tapasztalható forgalom jól leírható megbízható kísérleteken keresztül. Áttekintve a rejtett paraméterek meghatározására kidolgozott statisztikai becslési módszereket, a hálózati pufferek viselkedéséről számot adó tomográfias kiértékelő eljárásokat dolgoztunk ki. A bemutatott módszerek az ún. *expectation maximization* algoritmusok családjába tartoznak.

A dolgozatban felsorakoztatjuk a tomográfias mérés kiértékeléséből származó legfontosabb eredményeinket, melyek magukba foglalják az Európai Internet egy teljes napra rekonstruált forgalmi térképét, számot adnak a hálózat egyedi és összetett várakozási sorainak statisztikai jellemzőiről, feltárják a háttér forgalomban tapasztalható fraktál szerkezetet és segítségével bemutatunk egy, az Internetre jellemző, új tapasztalati skálázási törvényt.

A bevezetett becslési módszereinket megbízható mérési adatokra alkalmazhatjuk. Ennek biztosítására, nemzetközi összefogásban mérőhálózatot építettünk ki, az Etomic Hálózat mérési Infrastruktúrát. A mérési infrastruktúra Európa számos partner intézményben üzemeltetett eszköze hozzájárul a pontos tomográfias mérőprogramok futtatására. A mérőrendszer segítségével begyűjtött nyersadatokat Virtuális Obszervatóriumban publikáljuk, lehetővé téve az adatok más aspektusból történő jövőbeli elemzését, megteremtve annak feltételét, hogy más méréssel foglalkozó kutatási csoportok adataival együtt bonyolultabb kérdések megválaszolására alkalmasak legyenek.

A kísérletező munka mellett elméleti modellek kidolgozásával is foglalkozunk. Bemutatunk egy hálózati modellt, mely alkalmas számot adni a forgalmi torlódás jelenségére. A modell segítségével levezettük a dugulási hullámot hűen jellemző

annak alakjáról és terjedési sebességéről számot adó formulákat. Eredményeinket a paraméterek széles tartományán megvizsgáltuk hálózat szimulációs környezetek segítségével. A szimulációs nyersadatok alapján képleteinket teszteltük.

Hálózatkutatással kapcsolatos munkánk a dolgozat megírásával nem zárul le, a megválaszolatlan kérdések további kutatásaink irányát rajzolják elő.

8. fejezet

Summary

In the dissertation the traffic characteristics of computer communication networks has been studied with the attitude of both an experimental and a theoretical physicist. We have demonstrated that each agent's packet regulation dynamics, which rule the Internet overwhelmingly, are competing for the available resources in the network such a way that introduce interesting couplings. Due to the interactions of the agents strange congestion patterns evolve in the system.

Our goal was to found and distribute new measurement methods, by the help of which the traffic, present in the Internet, can be well tracked and described based on reliable experiments. After having overviewed the statistical methods elaborated to reveal hidden parameters of the systems, we developed new tomography procedures, which are capable of showing the queueing delay statistics of the internal buffers of the network. These algorithms belong to the family of the *expectation maximization* methods.

In the dissertation we enumerate the most important results of the queueing delay tomography methods. Our results involve the reconstruction of the congestion map of the European Internet. Network tomography also accounts for the statistical characteristics of delays over the individual and the compound links, the study of the former enables us to reveal the fractal structure of the background traffic. The detailed analysis of the average queueing delays show a new scaling property of the computer networks.

Our tomography approximation methods require reliable input data from measurements. In order to achieve this, we built a measurement facility in an international collaboration, which is called European Traffic Observatory Measurement Infrastructure, (*EtoMic*). The hardware components of this infrastructure, installed at various partner institutions, make it possible to run the precise measurement applications of the network tomography. All the raw experiment data collected in our measurements are stored and published in the databases of the Network Measurement Virtual Observatory, to enable the study of the data from different aspects in the future. Theoretically this approach also provides the means to run joint complex data analysis with outside network measurement data

sources.

Besides the experimental work we are also working on new theoretical models of the network. We presented a network model, which grasps the congestion phenomenon found in the network traffic. With the help of this model we were able to derive formulae describing the spatial envelope of the congestion pattern and its speed. The results of the model were tested against network simulation data over a wide range of the parameters.

Our contribution to science of networks is continued on after this work. The numerous unanswered questions arisen so far motivate our work in the future.

Köszönetnyilvánítás

Mindenek előtt köszönettel tartozom témavezetőmnek Vattay Gábornak, aki bevezetett a kommunikációs hálózatok forgalomvizsgálatának és -modellezésének tárgykörébe. Azon túl, hogy kihívásokkal tarkított munkát teremtett nekem és kollegáimnak, segítőkészsége végig nagyon sokat jelentett a dolgozat alapjául szolgáló kutatómunka elvégzésében.

Szeretném megköszönni a Komplex Rendszerek Fizikája Tanszék, a Collegium Budapest, Institute for Advanced Study és a Kommunikációs Hálózatok Laboratórium vezetőinek és valamennyi munkatársának, hogy munkámhoz megteremtették a biztonságos, nyugodt, barátságos környezetet.

Külön köszönettel tartozom munkatársaimnak és szerzőtársaimnak az együtt elért eredményekért, gondolatserkentő segítségükért. Csabai István, Simon Gábor, Pollner Péter, Vaderna Péter, Hága Péter, Mátray Péter, Hullár Béla, Laki Sándor, Purger Norbert, Dobos László, Éltető Tamás, Fekete Attila, Ulisses Alonso, Eduardo Magana, Daniel Morato, Javier Aracil, Anat Hapern, Ohad Serfaty és Yuval Shavitt sokat segített munkám során.

Köszönettel tartozom a Nemzeti Kutatási és Fejlesztési Program (*NKFP 02-032-04*), valamint (*OMFB-00635/2005*), és az európai uniós IST FET COMPLEXITY EVERGROW Integrated Project, Monitoring and Measurement in the Next Generation Technologies pályázatok anyagi támogatásáért.

Köszönettel tartozom feleségemnek és kollegáimnak a kézirat gondos átolvasásáért, a hibák észrevételéért és építő jellegű gondolataikért.

Legnagyobb köszönettel családomnak és barátaimnak tartozom, akik végig lelkesítőleg támogattak a dolgozatírás közben.

Irodalomjegyzék

- [1] M. Waldrop. Darpa and the internet revolution. http://www.darpa.mil/Docs/Internet_Development_200807180909255.pdf, 2008.
- [2] Y. Hyun. Walrus - graph visualization tool. <http://www.caida.org/tools/visualization/walrus/>, 2005.
- [3] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, 2002.
- [4] L. Kocarev and G. Vattay. *Complex Dynamics in Communication Networks*. Springer, 2005.
- [5] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *ACM SIGCOMM*, Boston, 1999.
- [6] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. *Comput. Commun. Rev.*, 29:251–262, 1999.
- [7] R. Govindan and H. Tangmunarunkit. Heuristics for internet map discovery. In *IEEE Infocom*, March 2000.
- [8] A. Vazquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of the internet. *Phys. Rev. E*, 65(6):066130, 2002.
- [9] R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: a statistical physics approach*. Cambridge University Press, 2004.
- [10] R. Albert, H. Jeong, and A. L. Barabási. Error and attack tolerance of complex networks. *Nature*, 406:378–382, 2000.
- [11] G. Palla, I. Derényi, I. Farkas, and T. Vicsek. Statistical mechanics of topological phase transitions in networks. *Phys. Rev. E*, 69:046117, 2004.
- [12] M. A. de Menezes and A. L. Barabási. Fluctuations in network dynamics. *Phys. Rev. Lett.*, 92:028701, 2004.
- [13] B. Tadić and G. J. Rogers. Packet transport on scale free networks. *Advances in Complex Systems*, 5:445–456, 2002.

- [14] B. Tadić, S. Thurner, and G. J. Rogers. Traffic on complex networks: Towards understanding global statistical properties from microscopic density fluctuations. *Phys. Rev. E*, 69(3):036102, 2004.
- [15] S. Sreenivasan, R. Cohen, E. López, Z. Toroczkai, and H. E. Stanley. Communication bottlenecks in scale-free networks. *CoRR*, abs/cs/0604023, 2006.
- [16] W. E. Leland, M. S. Taqq, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic. In *ACM SIGCOMM*, pages 183–193, San Francisco, California, 1993.
- [17] W. E. Leland, M. S. Taqq, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, pages 1–15, 1994.
- [18] A. Veres and M. Boda. The chaotic nature of tcp congestion control. In *IEEE INFOCOM 2000*, 2000.
- [19] I. Csabai. 1/f noise in computer network traffic. *J. Phys. A: Math. Gen.*, 27(12):417–421, 1994.
- [20] M. Takayasu, H. Takayasu, and T. Sato. Critical behaviors and 1/f noise in information traffic. *Phys. A*, 233:824–834, 1996.
- [21] L. Kleinrock. *Queueing Systems*, volume 1–2. Theory, Wiley Interscience, New York, 1975–1976.
- [22] V. Jacobson. Congestion avoidance and control. In *ACM SIGCOMM*, pages 314–329, 1988.
- [23] V. Jacobson. Modified TCP congestion avoidance algorithm. Technical report, end2end-interest mailing list, April 1990.
- [24] S. Floyd and V. Paxson. Difficulties in simulating the internet. *IEEE/ACM Transactions on Networking*, 9:392–403, 2001.
- [25] S. McCanne and S. Floyd. UCB/LBNL/VINT network simulator – ns (version 2). <http://www-mash.cs.berkeley.edu/ns/>, 1995.
- [26] The Cooperative Association for Internet Data Analysis. Caida. <http://www.caida.org/>.
- [27] National Laboratory for Applied Network Research. Nlanr. <http://www.nlanr.net>, 2006.
- [28] Internet Measurement Data Catalog. Datcat. <http://www.datcat.org/>.

- [29] M. Allman and V. Paxson. Issues and etiquette concerning use of shared measurement data. In *ACM SIGCOMM/USENIX Internet Measurement Conference*, October 2007.
- [30] R. Pang, M. Allman, V. Paxson, and J. Lee. The devil and packet trace anonymization. *ACM Computer Communication Rev.*, 36, January 2006.
- [31] M. Roughan and Y. Zhang. Secure distributed data mining and its application in large-scale network measurements. *SIGCOMM Computer Communication Rev.*, 36, 2006.
- [32] K. Claffy. Internet measurement and data analysis: topology, workload, performance and routing statistics. In *NAE 1999*, 1999.
- [33] A. Pásztor. *Accurate active measurement In the internet and its applications*. Phd thesis, The University of Melbourne, 2003.
- [34] M. Crovella and B. Krishnamurthy. *Internet Measurement: Infrastructure, Traffic and Applications*. John Wiley and Sons Inc., 2006.
- [35] Dante. Geant2 tools portal. <http://stats.geant2.net/>.
- [36] B. Augustin, X. Cuvelier, B. Orgogozo, F. Viger, T. Friedman, M. Latapy, C. Magnien, and R. Teixeira. Avoiding traceroute anomalies with paris traceroute. In *6th ACM SIGCOMM*, 2006.
- [37] P. Hágá. *Statistikai fizikai módszerek az Internet forgalom aktív mérésében*. doktori értekezés, ELTE TTK Fizika Doktori Iskola, 2007.
- [38] Ever Growing Complex Networks. Evergrow homepage. <http://www.evergrow.org/>.
- [39] European Traffic Observatory Infrastructure. Etomic homepage. <http://www.etomic.org/>.
- [40] D. Morató, E. Magaña, M. Izal, J. Aracil, F. Naranjo, F. Astiz, U. Alonso, I. Csabai, P. Hágá, G. Simon, J. Stéger, and G. Vattay. The european traffic observatory measurement infrastructure (etomic): A testbed for universal active and passive measurements. In *Tridentcom 2005*, pages 283–289, Trento, Italy, 23–25 February 2005.
- [41] E. Magaña, D. Morató, M. Izal, J. Aracil, F. Naranjo, F. Astiz, U. Alonso, I. Csabai, P. Hágá, G. Simon, J. Stéger, and G. Vattay. The european traffic observatory measurement infrastructure (etomic). In *IPOM 2004*, pages 165–169, Bijing, China, 11–13 Octobre 2004.
- [42] Endace homepage. <http://www.endace.com>.

- [43] PlanetLab. An open platform for developing, deploying and accessing planetary scale services. <http://www.planet-lab.org/>, 2003–.
- [44] The Dimes project. <http://www.netdimes.org/>.
- [45] HGP. Human genome project. http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml, –2003.
- [46] D. G. York and et al. The sloan digital sky survey: Technical summary. *The Astronomical Journal*, 120:1579–1587, 2000.
- [47] A. S. Szalay, T. Budavári, T. Mali, J. Gray, and A. Thakar. Web services for the virtual observatory. In *SPIE Conference on Advanced Telescope Technologies*, volume 4846, 2002.
- [48] P. A. Gutierrez, A. Bulanza, M. Dabrowski, B. Koskina, J. Quittek, C. Schmoll, F. Strohmeier, A. Vidács, and S. Zs. Kardos. A scalable system for sharing internet measurements. In *IPS-MoMe 2005 Workshop*, Warsaw, Poland, 2005.
- [49] S. Niccolini, S. Tartarelli, J. Quittek, and M. Swany. How to store trace-route measurements and related metrics. <http://tools.ietf.org/html/draft-niccolini-ippm-storetraceroutes-03>.
- [50] K. Cho, K. Mitsuya, and A. Kato. Traffic data repository at the wide project. In *USENIX 2000, FREENIX Track*, San Diego, CA, June 2000.
- [51] Moment. Monitoring and measurement in the next generation technologies. <http://www.fp7-moment.eu/>.
- [52] Y. Vardi. Network tomography: estimating source-destination traffic intensities from link data. *American Statistical Association*, 91:365, 1996.
- [53] M. Coates and R. Nowak. Network tomography for internal delay estimation. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, May 2001.
- [54] N. Duffield, J. Horowitz, F. L. Presti, and D. Towsley. *Network Delay Tomography from End-to-end Unicast Measurements*. Springer Verlag - Lecture Notes in Computer Science, Berlin, 2001.
- [55] M. Coates and R. Nowak. Network loss inference using unicast end-to-end measurement. In *ITC Conf. IP Traffic, Modelling and Management*, Monterey, CA, September 2000.
- [56] N. Duffield, F. L. Presti, V. Paxson, and D. Towsley. Inferring link loss using striped unicast probes. In *IEEE Infocom 2001*, Anchorage, AK, April 2001.

- [57] R. Cáceres, N. G. Duffield, S. B. Moon, and D. Towsley. Inference of internal loss rates in the mbone. In *IEEE/ISOC Global Internet 1999*, December 1999.
- [58] R. Cáceres, N. G. Duffield, J. Horowitz, D. Towsley, and T. Bu. Multicast-based inference of network-internal characteristics: Accuracy of packet loss estimation. In *IEEE Infocom 1999*, New York, March 1999.
- [59] N. G. Duffield and F. L. Presti. Multicast inference of packet delay variance at interior network links. In *IEEE Infocom 2000*, Tel Aviv, Israel, March 2000.
- [60] F. L. Presti, N. Duffield, J. Horowitz, and D. Towsley. Multicast-based inference of network-internal delay distributions. In *IEEE/ACM Transactions on Networking*, December 2002.
- [61] T. Bu, F. L. Presti, N. Duffield, and D. Towsley. Network tomography on general topologies. In *ACM Sigmetrics 2002*, Marina del Rey, CA, 2002.
- [62] M. Rabbat, R. Nowak, and M. J. Coates. Multiple source, multiple destination network tomography. In *IEEE/ACM Transactions on Networking*, December 2002.
- [63] M. Rabbat, R. Nowak, and M. J. Coates. Multiple source, multiple destination network tomography. In *IEEE Infocom*, Hong Kong, March 2004.
- [64] S. Moon, P. Skelly, and D. Towsley. Estimation and removal of clock skew from network delay measurements. In *Infocom 1999*, New York, March 1999.
- [65] V. Paxson. On calibrating measurements of packet transit times. In *SIGMETRICS*, June 1998.
- [66] G. J. McLachlan and T. Krishnan. *The EM algorithm and extensions*. John Wiley, New York, 1997.
- [67] C. L. Lawson and B. J. Hanson. *Solving Least Squares Problems*. Prentice-Hall, Englewood Cliffs, NJ, 1974.
- [68] GÉANT homepage. The pan-european data communications network for research and education. <http://www.geant.net/>, 2004–2005.
- [69] G. Liang and B. Yu. Maximum pseudo-likelihood estimation in network tomography. In *IEEE Transactions on Signal Processing*, volume 51, 2003.
- [70] B. J. Frey. *Graphical models for machine learning and digital communication*. The MIT Press, Cambridge, Mass, 1998.

- [71] I. Norros. On the use of fractional brownian motion in the theory of connectionless networks. *IEEE Journal on Selected Areas in Communications*, 16:953–962, August 1995.
- [72] K. Papagiannaki, S. Moon, C. Fraleigh, P. Thiran, F. Tobagi, and C. Diot. Analysis of measured single-hop delay from an operational backbone network. In *INFOCOM*, 2000.
- [73] M. Izal and J. Aracil. On the influence of self-similarity on optical burst switching traffic. In *GLOBECOM*, volume 3, pages 2308–2312.
- [74] I. Antoniou, V. V. Ivanov, and Y. L. Kalinovsky. Kinetic model of network traffic. *Phys. A*, 308, 2002.
- [75] D. Chowdhury, L. Santen, and A. Schadschneider. Statistical physics of vehicular traffic and some related systems. *Phys. Rep.*, 329:199–329, 2000.
- [76] B. S. Kerner and P. Konhäuser. *Phys. Rev. E*, 50:54, 1994.
- [77] K. Nagel and M. Paczuski. Emergent traffic jams. *Phys. Rev. E*, 51:2909, 1995.
- [78] O. Biham, A. Middleton, and D. Levine. Self-organization and a dynamical transition in traffic-flow models. *Phys. Rev. A*, 46:6124, 1992.
- [79] G. Simon and I. Csabai. The analogies of highway and computer network traffic. *Phys. A*, 307:516–526, 2002.
- [80] T. Huisinga, R. Barlovic, W. Knospe, A. Schadschneider, and M. Schreckenberg. A microscopic model for packet transport in the internet. *Phys. A*, 294:249, 2001.
- [81] T. Musha and H. Higuchi. 1/f fluctuation of a traffic current on an expressway. *Jpn. J. Appl. Phys.*, 15:1271–1275, 1976.
- [82] K. Nagel and M. Schreckenberg. A cellular automaton model for freeway traffic. *J. Phys. I*, 2:2221, 1992.
- [83] K. Nagel and H. J. Hemann. Deterministic models for traffic jams. 254, 1993.
- [84] M. Leibig. Pattern-formation characteristics of interacting kinematic waves. *Phys. Rev. E*, 49:184, 1994.
- [85] K. Fukuda, H. Takayasu, and M. Takayasu. Spatial and temporal behaviour of congestion in internet traffic. *Fractals*, 7:23–31, 1999.

- [86] W. Gong, Y. Liu, V. Misra, and D. Towsley. On the tails of web file size distributions". In *Annual Allerton Conference on Communication, Control, and Computing*, 2001.